Titre: Title:	Architecture de survivabilité aux pannes catastrophiques pour réseaux basés sur MPLS
Auteur: Author:	Antoine Lemay
Date:	2005
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Lemay, A. (2005). Architecture de survivabilité aux pannes catastrophiques pour réseaux basés sur MPLS [Master's thesis, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/8388/

UNIVERSITÉ DE MONTRÉAL

ARCHITECTURE DE SURVIVABILITÉ AUX PANNES CATASTROPHIQUES POUR RÉSEAUX BASÉS SUR MPLS

ANTOINE LEMAY
DÉPARTEMENT DE GÉNIE INFORMATIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE INFORMATIQUE)
AVRIL 2005

Your file Votre référence
ISBN: 978-0-494-47674-1
Our file Notre référence
ISBN: 978-0-494-47674-1

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

ARCHITECTURE DE SURVIVABILITÉ AUX PANNES CATASTROPHIQUES POUR RÉSEAUX BASÉS SUR MPLS

Présenté par : LEMAY Antoine

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. CHAMBERLAND Steven, Ph.D., président

M. PIERRE Samuel, Ph.D., membre et directeur de recherche

M. FERNANDEZ José, Ph.D., membre

REMERCIEMENTS

Je tiens à remercier en premier lieu mon directeur de recherche, monsieur Samuel Pierre, pour son appui constant et ses conseils avisés dont j'ai bénéficié grandement au cours de mon travail.

Je tiens également à remercier monsieur Yves Lemieux de Ericsson Recherche Canada pour son dévouement dans le cadre de la Chaire CRSNG-ERICSSON en Systèmes Réseautiques Mobiles de Prochaines Générations.

Il me faut aussi remercier chaleureusement mes confrères et consoeurs du LARIM, plus particulièrement mes collèges de bureau Mélissa Georges, Nabil Harrabida et Méral Shirazipour.

Finalement, j'aimerais offrir ces derniers remerciements à ma famille et ma copine qui m'ont soutenu tout au long de ce travail.

RÉSUMÉ

La gestion des catastrophes dans le domaine des télécommunications a toujours été un peu négligée. Seuls les réseaux militaires se sentaient concernés par ce type de pannes et seuls les militaires avaient les moyens de déployer des mesures de protection. Avec les attentats du 11 septembre 2001, une idée de l'impact de la perte d'un réseau de télécommunications s'est propagée dans la communauté. Compte tenu de l'envergure catastrophique de ces pannes, il a été jugé inadmissible de permettre ce genre d'événement. L'intérêt pour la survivabilité des réseaux d'information aux pannes catastrophiques a donc pris son essor.

Toutefois, les méthodes traditionnelles de protection contre les pannes ne sont pas adéquates pour les réseaux de prochaine génération. En effet, ces méthodes sont extrêmement rapides pour répondre aux exigences de qualité de service des réseaux commutés, mais elles ont été conçues pour des pannes simples. Les méthodes traditionnelles pour les réseaux de paquets sont très résistantes aux pannes multiples, mais sont trop lentes pour le trafic nécessitant une qualité de service. Finalement, les méthodes militaires se basent sur la duplication à outrance et sont trop onéreuses pour avoir une applicabilité appréciable.

Pour pallier ces lacunes, ce mémoire propose deux méthodes de gestion des pannes catastrophiques pour les réseaux métropolitains. Ces méthodes s'inspirent de la littérature pour fournir deux nouvelles approches basées sur les architectures IP et MPLS. Ces approches sont comparées à la méthode de commutation de protection en aval par MPLS en terme de délai et de quantité de trafic restauré. Après comparaison, on constate que la solution avec division de bande passante offre une augmentation de la prédictibilité malgré une dégradation du délai, et que la solution sans division de bande passante offre une augmentation de la quantité de trafic restauré en sacrifiant légèrement sur le délai de recouvrement.

Ces constatations sont soutenues par une preuve de fonctionnement de la solution avec division de bande passante, réalisée à l'aide du simulateur OPNET. Cette expérience montre qu'il est possible d'obtenir une quantité de trafic restaurée

équivalente à celle restaurée en utilisant la commutation de protection par MPLS. Puisque la solution sans division de la bande passante est démontrée plus performante que la solution implémentée, nous pouvons conclure qu'elle peut dépasser les performances de MPLS. Cette preuve de fonctionnement est complémentée par une analyse, à l'aide de MATLAB, de l'influence des différents paramètres sur nos solutions. On montre que dans toutes les circonstances, il est possible d'égaler ou de surpasser la performance de MPLS en termes de quantité de trafic restaurée. Finalement, nous présentons une étude statistique qui montre la validité des résultats générés par le modèle analytique et qui montre aussi le caractère de prédictibilité de la solution 1.

ABSTRACT

The management of disasters in the field of telecommunications has always been somewhat neglected. Only military networks were concerned about this type of failures and only military organizations had the financial means to deploy countermeasures. With the attacks of September 11th 2001, the idea of the consequences of loosing a telecommunication network has spread in the community. Because of the gravity of the consequences of a catastrophic failure, letting such a disaster occur is not an option. Thus, there has been a great rise of the interest in the field of survivability of critical information systems against catastrophic failures.

However, the traditional methods to protect networks against failures are not adapted to next generation networks. The methods traditionally used in circuit-switched networks are very fast (to satisfy the quality of service requirements), but are designed for single failures. The traditional methods of packet-switched networks are very resilient to catastrophic failures, but are too slow to satisfy the quality of service requirements for sensitive traffic. Lastly, the methods used by the various military organizations are based on excessive duplication and are too costly to be applicable on a large scale.

To solve these problems, this research proposes two methods of handling the management of catastrophic failures for metro area networks. These methods are inspired from existing methods in the literature and are based on the IP and MPLS architectures. The proposed methods are compared to MPLS protection switching at the ingress node in terms of delay and amount of traffic restored. After comparison, we see that the approach with bandwidth division offers an increase in predictability at the expense of delay and that the approach without bandwidth division offers an increase in the amount of traffic restored at the expense of time to repair.

These findings are supported by a proof of concept for the solution with bandwidth division mad with the OPNET simulator. This experiment shows that it is possible to obtain an amount of restored traffic similar to the amount restored by MPLS

protection switching. Since the solution with no bandwidth division is shown to have a better performance, we can conclude that this second solution can surpass the performance of MPLS protection switching. The proof of concept is complemented by a numerical analysis of the influence of the various parameters on our solutions with the help of the MATLAB software. We show that, in all circumstances, it is possible to equal or improve on the performance of MPLS protection switching in terms of amount of traffic restored. Finally, we present a statistical analysis showing the validity of the results generated from the analytical model. That study also shows increased predictability for the solution with bandwidth division.

TABLE DES MATIÈRES

REMERCIEMENTS iv
RÉSUMÉ v
ABSTRACT vii
TABLE DES MATIÈRES ix
LISTE DES FIGURES xii
LISTES DES TABLEAUX xiv
LISTE DES SIGLES ET DES ABBRÉVIATIONS xv
LISTE DES ANNEXES xvii

CHAPITRE I : INTRODUCTION 1
1.1 Définitions et concepts de base 1
1.2 Éléments de la problématique 4
1.3 Objectifs de recherche 6
1.4 Plan du mémoire 7

CHAPITRE II : ANALYSE DES MÉCANISMES DE SURVIVABILITÉ 8
2.1 Particularités du champ d'étude de la survivabilité 8
2.1.1 Définition traditionnelle de panne 8
2.1.2 Ajouts au modèle traditionnel de la panne 10
2.1.3 Modélisation mathématique de la survivabilité 11
2.1.4 Coordination des mécanismes de survivabilité 13
2.2 Survivabilité de bas niveau 13
2.2.1 Topologie en anneau 14
2.2.2 Topologie en maille 15
2.2.3 Topologies hybrides 17
2.2.4 Survivabilité SONET 18

2.3 Survivabilité de niveau intermédiaire 20
2.3.1 Survivabilité IP 20
2.3.2 Survivabilité MPLS 20
2.3.3 Protection par commutation 22
2.3.4 Protection par reroutage rapide 23
2.3.5 Modèle de RD-QoS 25
2.4 Survivabilité de haut niveau 26
2.4.1 Particularités de la survivabilité de haut niveau 27
2.4.2 Architectures réparties 28
2.4.3 Réseaux actifs 29
2.4.4 Applications reconfigurables 31
2.5 Conclusion 33

CHAPITRE III : MÉCANISMES DE SURVIVABILITÉ
POUR FAILLE MAJEURE BASÉS SUR MPLS 34
3.1 Principes et fondements de la solution proposée 34
3.1.1 Principes 34
3.1.2 Fondements de l'approche 37
3.2 Étude du délai 40
3.2.1 Méthode de protection par commutation, classe C1 40
3.2.2 Méthode avec division de la bande passante, classe C1 42
3.2.3 Méthode sans division de la bande passante, classe C1 44
3.2.4 Étude du délai pour la classe C2 47
3.3 Étude de la survivabilité 48
3.3.1 Méthode de commutation à l'ingress, classe C1 48
3.3.2 Méthode avec division de la bande, classe C1 49
3.3.3 Méthode sans division de la bande, classe C1 49
3.3.4 Étude de la survivabilité pour la classe C2 51
3.4 Modélisation des architectures de protection 52

3.4.1 Caractéristiques communes des architectures de protection 52
3.4.2 Architecture avec division de la bande passante 55
3.4.3 Architecture sans division de la bande passante 57
3.5 Prototypage 58

CHAPITRE IV : ANALYSE DES RÉSULTATS 62
4.1 Plan d'expérience 62
4.1.1 Objectifs 62
4.1.2 Métriques 63
4.1.3 Tests optimaux 65
4.1.4 Tests réalisés 66
4.2 Résultats de simulation 67
4.2.1 Environnement de simulation OPNET 67
4.2.2 Modèle de simulation 71
4.3.3 Résultats 73
4.3 Modèle analytique 75
4.3.1 Modèle basé sur l'espérance mathématique 76
4.3.2 Modèle probabiliste 83

CHAPITRE V 93
CONCLUSION 93
5.1 Synthèse des travaux 93
5.2 Limitations des travaux 95
5.3 Indications de recherche future 96

BIBLIOGRAPHIE 99
ANNEXE 104

LISTE DES FIGURES

Figure 2.1 Panne de lien 9
Figure 2.2 Panne de nœud 9
Figure 2.3 Partitionnement réseau 10
Figure 2.4 Partitionnement d'un réseau administré centralement 11
Figure 2.5 Topologie en anneau 14
Figure 2.6 Réseau complètement maillé 16
Figure 2.7 Topologie en anneau augmenté 17
Figure 2.8 Topologie en cycle-p 18
Figure 2.9 Protection par commutation 22
Figure 2.10 Protection par reroutage rapide 24
Figure 2.11 Problème de cohérence des étiquettes 25
Figure 2.12 Architecture de restauration RAPTOR 32
Figure 3.1 Architecture générale de survivabilité 53
Figure 3.2 Mécanismes de survivabilité de la solution 1 56
Figure 3.3 Mécanismes de survivabilité de la solution 2 57
Figure 4.1 Trafic dans les LSPs avant la panne 73
Figure 4.2 Trafic dans les LSPs après la panne 74
Figure 4.3 Quantité de trafic restauré pour la solution 1, n = 2 77
Figure 4.4 Quantité de trafic restauré pour la solution 1, n = 3 78
Figure 4.5 Quantité de trafic restauré pour la solution 1, n = 5 79
Figure 4.6 Quantité de trafic restauré pour la solution 1, n = 10 79
Figure 4.7 Quantité de trafic restauré pour la solution 2, n = 2 80
Figure 4.8 Quantité de trafic restauré pour la solution 2, n = 3 81
Figure 4.9 Quantité de trafic restauré pour la solution 2, n = 5 82
Figure 4.10 Quantité de trafic restauré pour la solution 2, n = 10 82
Figure 4.11 Moyenne de trafic restauré pour la solution 1, n = 3 84

Figure 4.12 Moyenne de trafic restauré pour la solution 1, n = 5 84
Figure 4.13 Moyenne de trafic restauré pour la solution 1, n = 10 85
Figure 4.14 Variance du trafic restauré pour la solution 1, n = 3 86
Figure 4.15 Variance du trafic restauré pour la solution 1, n = 5 87
Figure 4.16 Variance du trafic restauré pour la solution 1, n = 10 87
Figure 4.17 Moyenne de trafic restauré pour la solution 2, n = 3 88
Figure 4.18 Moyenne de trafic restauré pour la solution 2, n = 5 89
Figure 4.19 Moyenne de trafic restauré pour la solution 2, n = 10 89
Figure 4.20 Variance du trafic restauré pour la solution 2, n = 3 90
Figure 4.21 Variance du trafic restauré pour la solution 2, n = 5 91
Figure 4.22 Variance du trafic restauré pour la solution 2, n = 10 91

LISTES DES TABLEAUX

Tableau 2.1 Classes de Résilience 26
Tableau 3.1 Comparaison des délais pour le trafic de type C1 46
Tableau 3.2 Compraison de la survivabilité du trafic de type C1 51

LISTE DES SIGLES ET DES ABBRÉVIATIONS

ADM	Add-Drop Multiplexer
APS	Automatic Protection Switching
ASP	Automatic Switch to Protection
ASW	Automatic Switch to Working
ATM	Asynchronous Transfer Mode
BLSR	Bidirectional Link Switched Ring
bps	bits per second
CDMA	Code Division Multiple Acces
CSPF	Constrained Shortest Path First
Diff-Serv	Differentiated Services architecture
DQ	Default Queue
FEC	Forwarding Equivalence Class
FIS	Fault Indication Signal
GMPLS	Generalised MPLS
GRP	Groupe à Risque Partagé
GSM	Global System for Mobile Communication
HTTP	HyperText Transfer Protocol
IP	Internet Protocol
Kbps	Kilobits per second
LLQ	Low Latency Queue
LSP	Label Switched Path
LSR	Label Switching Router
MAN	Metropolitan Area Network
MPLS	MultiProtocol Label Switching
MPλS	MultiProtocol wavelenght Switching
OSI	Open Systems Interconnection

OSPF	Open Shortest Path First
PLR	Point of Local Repair
PML	Point for Merging Labels
QoS	Quality of Service
RD-QoS	Resilisence Differentiated Quality of Service
RSVP	Ressource Reservation Protocol
RSVP-TE	Traffic Engineering extension for RSVP
SONET	Synchronous Optical NETwork
UMTS	Universal Mobile Telecommunication System
UPSR	Unidirectional Path Switched Ring
VC	Virtual Circuit
WiFi	Wireless Fidelity

LISTE DES ANNEXES

Annexe 1 : Architecture de protection SONET 104

Annexe 2 : Architecture de protection IP 105

Annexe 3 : Topologie du réseau de test 106

CHAPITRE I
INTRODUCTION

Le besoin de communiquer de l'être humain est connu depuis la nuit des temps. Alors que jadis les réseaux de communications se bornaient aux routes et aux pigeons voyageurs, aujourd'hui nous disposons de puissants réseaux de communications tel que l'Internet. Avec les développements technologiques, on voit de plus en plus apparaître la convergence de ces différents réseaux de communication. Cette convergence crée de nouveaux défis, notamment au niveau de la gestion de la mobilité, de la qualité de service et de la sécurité. Parmi ces défis, le problème de la qualité de service est d'une importance capitale pour les réseaux de nouvelles générations, intégrant voix et vidéo sur IP (Internet Protocol). Toutefois, il est impossible d'implémenter une qualité de service sur un réseau qui n'est pas fiable. Pour augmenter la fiabilité, il est courant de recourir à la survivabilité des réseaux, objet de ce présent mémoire qui étudie les différents mécanismes de survivabilité applicable pour les pannes catastrophiques. Ce chapitre d'introduction présentera premièrement quelques définitions et concepts de bases liés au champ d'étude traité. Ensuite, les différents éléments de la problématique seront développés en spécifiant l'intérêt du domaine d'étude et en listant les principales contraintes. Par la suite, les objectifs de recherche et les principaux résultats attendus seront définis. Enfin, le plan du mémoire sera esquissé.

1.1 Définitions et concepts de base

La survivabilité est définie dans [28] comme la capacité d'un système à remplir sa mission, en un temps raisonnable, en présence d'événements dommageables de tous types. Pour sa part, la notion de *mission* réfère à un groupe de requis de haut niveau ou de buts tandis que la notion de *temps raisonnable* réfère aux contraintes temporelles pour l'accomplissement de cette mission.

Dans le contexte d'un réseau de télécommunications, la survivabilité s'exprime par la capacité de livrer les services essentiels (et maintenir les propriétés essentielles de ces services comme la qualité de service) même si une portion significative du système cesse d'être disponible. Ce champ d'étude se distingue de la tolérance aux fautes qui, elle, se base sur la probabilité d'occurrence des fautes. Ainsi, les fautes peu probables seront ignorées. Par contre, dans une étude de survivabilité, ces fautes peu probables doivent être prises en compte si l'occurrence de celles-ci causait une défaillance inacceptable.

En se plaçant dans une perspective d'usager, la mesure de la survivabilité d'un réseau est la disponibilité des services en cas de pannes. Cette disponibilité comprend deux aspects : la connectivité et la qualité de service. La connectivité est la capacité de trouver un chemin (fonctionnel) entre l'usager et le nœud qu'il veut contacter dans le réseau. La qualité de service réfère plutôt à la qualité de la transmission sur ce chemin et l'effet sur différents paramètres comme le délai, la gigue et la perte de paquets. Ces paramètres influencent directement le fonctionnement de l'application tel que perçu par l'usager. Dans le cas d'IP, la qualité de service est fortement liée à la disponibilité de bande passante sur le réseau. Un service qui peut être rejoint par un chemin dans le réseau avec des critères de qualité de service adéquats est dit disponible. Un réseau assurant la disponibilité de ses services, même en cas de panne, est dit survivable.

Ainsi, pour obtenir un réseau survivable, on doit disposer d'un chemin entre le nœud source et le nœud destination. De plus, pour assurer la qualité de service, ce chemin doit disposer d'une bande passante supérieure ou égale à la bande passante nécessaire pour assurer les paramètres de qualité de services requis par le trafic en entrée. La ressource qui nous intéresse est donc l'élément de chemin. Dans un réseau de télécommunications ceci se traduit par la bande passante disponible sur les liens et la capacité de traitement des nœuds. La survivabilité d'un réseau est limitée par la disponibilité de ces éléments de chemins en cas de panne. La manipulation de ces éléments de chemin nous fournit deux outils pour implémenter la survivabilité : la redondance et la diversité.

La redondance est une duplication des ressources en parallèle au système courant. Le niveau de redondance est donné par le rapport entre les unités redondantes sur les unités fonctionnelles. Dans le cas où le système courant ne fonctionnerait pas, il est possible de prendre le système redondant (ou système de secours.) Ainsi, le système peut survivre à une panne dans le réseau. Pour s'assurer de la disponibilité des ressources de secours en cas de besoin, celles-ci doivent être inutilisées au moment de la panne. Les ressources de secours sont donc généralement laissées en attente. La redondance est donc peu efficace au niveau de l'utilisation des ressources mais permet de rétablir le trafic jusqu'au niveau de redondance du système. Dans un réseau de communications, lorsqu'on ajoute des nœuds ou des liens superflus, on augmente la redondance.

La diversité est la disposition des ressources de façon à minimiser les impacts de la défaillance d'une d'entre elles sur les autres ressources. La possibilité d'un téléphone cellulaire de communiquer par WiFi (Wireless Fidelity) pour contrer les pannes dans l'interface CDMA (Code Division Multiple Access) est un exemple de diversité. La diversité s'accompagne généralement de redondance puisqu'il faut habituellement ajouter de nouvelles ressources pour utiliser ces méthodes alternatives. Dans le cas d'un réseau de télécommunications, lorsqu'on augmente le nombre de chemins, on augmente la diversité. La diversité d'un réseau de télécommunications est intimement liée au degré des nœuds. Le degré d'un nœud est donné par le nombre de voisins auquel ce nœud est relié. Plus le degré est élevé, plus le nœud permet une grande diversité.

La diversité d'un réseau de télécommunications est plutôt liée à la *K-connectivité* du réseau (qui dépend en partie du degré des nœuds). Il existe deux types de K-connectivité : la *K-connectivité-arc* et la *K-connectivité-nœud.* On retrouve dans [25] la définition de chacun de ces concepts. Un graphe est dit K-connecté-arc si et seulement si chaque paire de nœuds est connectée par au moins K chemins disjoints d'arc. C'est-à-dire que chaque paire de nœuds est connectée par au moins K chemins tels qu'il n'existe aucun arc commun à deux de ces chemins. Un arc est dit K-connecté-nœud si et seulement si chaque paire de nœuds est connectée par au moins K chemins disjoints de nœud. C'est-à-dire que chaque paire de nœuds est connectée par au moins K chemins

tels qu'il n'existe aucun nœud commun à deux de ces chemins. Dans le cas d'une étude de survivabilité contre les pannes catastrophiques, pannes pour lesquelles il faut envisager les pannes de nœuds, il est nécessaire de considérer la K-connectivité-nœud, qui est plus contraignante que la K-connectivité-arc. Ainsi, la mesure de diversité dans un réseau sera donnée par la *connectivité nœud* du réseau qui est définie comme la plus grande valeur de K pour laquelle un réseau est K-connecté-nœud [25].

1.2 Éléments de la problématique

La disponibilité des réseaux de communications numériques s'est retrouvée à la base de la révolution de l'économie des grands pays industrialisés pour passer vers une économie du savoir. Aujourd'hui, une panoplie d'activités essentielles à l'économie, comme les transactions bancaires et le commerce électronique, se déroulent à l'aide de ces réseaux. Avec l'avènement des réseaux de prochaine génération, viendront s'ajouter plusieurs nouveaux services dont les appels téléphoniques et les services d'urgence 911. Il est clair que l'activité humaine devient de plus en plus dépendante de la disponibilité des réseaux de communications. D'ailleurs, le directeur de la cyber sécurité pour le département de la Sécurité Nationale Américaine, Amit Yoran, affirmait le 3 décembre 2003 : « Lorsqu'on observe la douzaine ou plus d'infrastructures critiques à propos desquelles le département du Homeland Security est particulièrement concerné, un fil directeur se dégage – leur dépendance par rapport à une infrastructure informatique robuste, fonctionnelle et sécure. »[1] Les attaques terroristes du 9 septembre ont mis à jour la fragilité de ce réseau à des attaques délibérées et ayant l'intention de causer des dommages. À partir de ce constat de fragilité, un consensus se dégage sur la nécessité d'assurer la survivabilité des réseaux supportant des infrastructures critiques à des attaques corrélées de grande envergure.

La résolution de ce problème est facile pour des réseaux de type IP. En effet, s'il n'y a pas de contraintes de qualité de service, il est uniquement nécessaire de se

[1] Amit Yoran, Director, National Cyber Security Division, National Cyber Security Summit, Santa Clara, California, December 3, 2003

préoccuper de la connectivité du réseau après une panne. Il est aisé de conserver la connectivité en ajoutant de la diversité. Cette connectivité n'est toutefois pas une mesure adéquate pour assurer la survivabilité des réseaux convergés intégrant la qualité de service. En effet, même si le chemin existe, il n'y a aucune garantie que ce lien soit capable de fournir la qualité de service requise à toutes les communications redirigées vers lui. Néanmoins, notons que, si le degré du réseau (c'est-à-dire le degré du nœud de plus faible degré) est suffisamment grand, il est possible d'assurer la survivabilité à des pannes multiples et corrélées.

Pour pallier ce problème de qualité de service, plusieurs méthodes ont été développées pour trouver deux chemins disjoints entre une source et une destination (entre autres [30], [31], [28]). Ainsi, advenant une panne sur un chemin, il serait possible d'emprunter le chemin de secours. En vérifiant les paramètres de qualité de service sur ce chemin de secours, il est possible d'assurer la qualité de service. Aussi, si les pannes sont indépendantes et ont une probabilité d'occurrence assez faible, il est possible de négliger l'occurrence de pannes simultanées sur le chemin principal et le chemin de secours. En effet, la probabilité de cet événement est le produit de la probabilité de panne sur chacun des chemins qui est, par hypothèse, petite. Par contre, cette hypothèse n'est pas valide dans le cas où les pannes seraient corrélées, puisque la probabilité de pannes simultanées ne correspond pas au produit des probabilités de pannes individuelles.

Une autre approche pour résoudre le problème d'incompatibilité entre la connectivité et la qualité de service est le surdimensionnement. Cette approche consiste à dimensionner les éléments de réseaux avec une taille suffisante pour accommoder à la fois le trafic normal et le trafic supplémentaire résultant de pannes en augmentant la redondance du réseau. Bien qu'il soit possible d'assurer la qualité de service en déployant suffisamment de matériel, l'investissement nécessaire pour garantir une qualité de service adéquate est prohibitif. Ce coût est d'autant plus important lorsqu'on tient compte de la possibilité de pannes multiples et corrélées. Cette approche, extrêmement coûteuse, est généralement uniquement à la portée des militaires.

On voit que les méthodes traditionnelles d'aborder le problème de la survivabilité ne permettent pas de résoudre le problème de pannes multiples et corrélées en intégrant la qualité de service à un coût raisonnable. Étant donné la dépendance accrue sur les réseaux d'information et la nécessité d'assurer la disponibilité des réseaux supportant les infrastructures critiques, ce mémoire s'inscrit dans le cadre de travaux de recherche visant à développer des mécanismes de survivabilité pour les réseaux d'information transportant des données de type mission critique.

1.3 Objectifs de recherche

L'objectif principal de ce mémoire est de concevoir une architecture permettant d'assurer la survivabilité des réseaux transportant de l'information de type mission critique pour du trafic conversationnel et meilleur effort. Cet objectif doit s'accomplir en assurant un temps de restauration inférieur aux limites imposées par la qualité de service, soit 140 ms pour du trafic conversationnel et quelques minutes pour du trafic meilleur effort. Plus spécifiquement ce mémoire vise à :

- Analyser les mécanismes de survivabilité existants en vue de déterminer leurs faiblesses relatives à la survivabilité à des pannes corrélées dans un domaine intégrant la qualité de service et de tirer des conclusions relatives aux règles gouvernant l'implémentation des mécanismes de survivabilité ;
- Concevoir une architecture qui sera basée sur la qualité de service des différentes classes de trafic et sur les ressources disponibles au réseau ;
- Évaluer l'impact de cette architecture sur la qualité de service associée aux différentes classes de service ;
- Comparer la performance de cette architecture par rapport à l'architecture MPLS (Multi-Protocol Label Switching) implémentée sur OPNET en temps que mécanisme assurant la meilleure survivabilité en respectant les contraintes de qualité de service.

Cette architecture sera développée pour un réseau d'accès métropolitain (MAN) ayant en entrée du trafic UMTS.

1.4 Plan du mémoire

Ce mémoire comporte quatre chapitres outre cette introduction qui en constitue le premier. Le second chapitre analyse les différents mécanismes de survivabilité les plus communs dans la littérature scientifique. Le troisième chapitre expose les mécanismes de survivabilité de l'architecture proposée ainsi que leur implémentation. Le quatrième chapitre détaille les expériences et les mesures effectuées pour l'analyse de performance et analyse les résultats en les comparants aux résultats obtenus à partir de l'architecture MPLS. Finalement, le cinquième et dernier chapitre fait une synthèse de résultats obtenus, détaille les limitations des travaux et propose des pistes de recherches futures.

CHAPITRE II
ANALYSE DES MÉCANISMES DE SURVIVABILITÉ

L'effondrement du secteur économique des télécommunications a précipité au premier plan l'ingénierie du trafic. Celle-ci a permis d'opérer plus efficacement les réseaux sans nécessiter d'investissement additionnel. De nos jours, on remarque un intérêt accru pour les applications relatives à la sécurité. Ainsi, la survivabilité, domaine de l'ingénierie du trafic étudiant la performance en cas de panne, est un devenu sujet de recherche populaire. Dans ce chapitre, les différents mécanismes de survivabilité que l'on peut retrouver dans la littérature seront exposés. Tout d'abord, nous présenterons les particularités du champ d'étude de la survivabilité, ensuite nous poursuivrons en détaillant les mécanismes de survivabilité existants pour les couches les plus basses du modèle OSI, les couches intermédiaires et les couches les plus hautes. Finalement, nous tirerons quelques conclusions relatives à l'implémentation des mécanismes de survivabilité.

2.1 Particularités du champ d'étude de la survivabilité

Puisque la survivabilité est reliée à la performance en présence d'une panne, les domaines de la survivabilité et de la tolérance aux fautes sont souvent confondus. On retrouve cependant des différences marquées, particulièrement au niveau de la définition traditionnelle de panne, des ajouts à ce modèle traditionnel de panne et de la modélisation du problème à résoudre. De plus, le problème additionnel de la coordination verticale se présente.

2.1.1 Définition traditionnelle de panne

La définition du dictionnaire de « panne » nous est donné comme suit : « Arrêt de fonctionnement accidentel et momentané. » Cette définition s'applique mal à des systèmes complexes comme les réseaux de communications puisque cet arrêt peut se

retrouver à n'importe quel niveau d'abstraction. Le concept de « Groupe à risque partagé » (ou GRP) a été introduit pour clarifier cette notion.

Un GRP est défini comme étant un ensemble d'éléments de réseau qui sont affectés simultanément par une faute spécifique ou un type spécifique de faute [22]. Par exemple, tous les passagers d'une voiture sont affectés par une faute de type « accident de voiture.» Il est donc évident que, lorsqu'on GRP subit une panne, les pannes des différents éléments constituant le GRP sont fortement corrélées. Par contre, il y a très peu de corrélation entre les pannes de GRPs distincts. Il est alors possible de regrouper tous les éléments d'un GRP en un « domaine de risque ». Ce domaine s'apparente à un domaine administratif et rend plus facile le traitement de fautes fortement corrélées.

La définition traditionnelle de panne prend en compte deux types de pannes : les pannes de lien et les panne de nœuds. Une panne de lien survient lorsque le lien connectant deux nœuds cesse de fonctionner, comme l'illustre la Figure 2.1. Dans cette situation, seulement le lien est inutilisable, les nœuds sont toujours fonctionnels.

Figure 2.1 Panne de lien

Une panne de nœud survient lorsqu'un nœud cesse de fonctionner normalement. On se retrouve alors dans la situation illustrée à la Figure 2.2. Dans ce cas, puisque le nœud ne peut plus effectuer de traitement, à la fois le nœud et les liens connectés à ce nœud sont inutilisables.

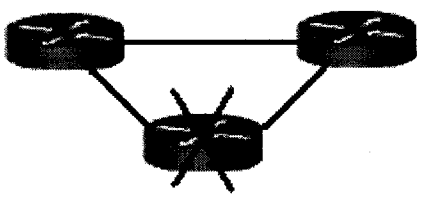

Figure 2.2 Panne de nœud

2.1.2 Ajouts au modèle traditionnel de la panne

Ces deux types de pannes correspondent à des pannes simples, c'est-à-dire des pannes n'affectant qu'un seul élément du réseau. Il est aussi possible de subir des pannes multiples. Si ces pannes ont toutes la même cause, on dit qu'on fait face à des pannes corrélées. La corrélation entre ces pannes émane du fait que ces éléments font partie du même GRP. Quelques exemples de pannes corrélées sont la panne de plan de contrôle, les attaques de type déni de service par inondation, les cas de congestion et les attaques physique coordonnées par des entités hostiles ou les forces de la nature. Lorsqu'on traite une panne corrélée, il est impossible de supposer l'indépendance statistique des pannes au niveau des éléments de réseau. Par le fait même, il est impossible de supposer qu'un chemin de secours pré-établi sera disponible en cas de panne.

Les pannes corrélées posent de nouveaux problèmes par rapport aux pannes simples. En particulier, on remarque l'apparition de problèmes comme le partitionnement et la non-masquabilité.

Bien que possible dans un scénario de panne simple, dans la plupart des cas, le partitionnement est un problème particulier aux pannes multiples. Il survient lorsqu'une panne divise le réseau en deux ou plusieurs parties non connexes. La Figure 2.3 illustre un partitionnement.

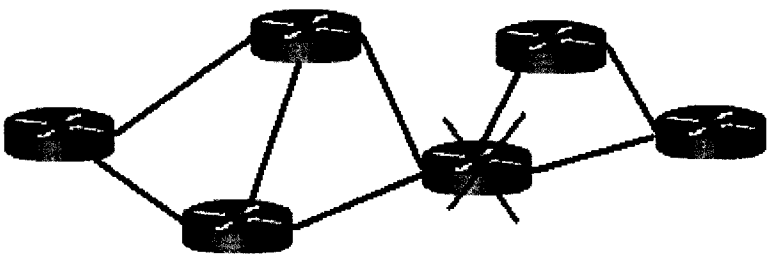

Figure 2.3 Partitionnement de réseau

On remarque que le résultat est deux composantes connexes, une contenant 3 nœuds à gauche et une autre contenant 2 nœuds à droite. Dans un réseau administré centralement, les composantes ne faisant pas partie de la composante connexe

comportant la console d'administration ne sont plus administrables. Elles ne peuvent donc plus fonctionner correctement. On obtient alors le scénario de panne illustré à la Figure 2.4.

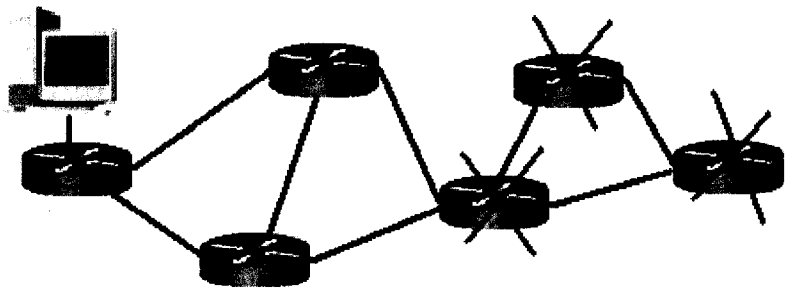

Figure 2.4 Partitionnement d'un réseau administré centralement

Similairement, toute connexion entre des nœuds de composantes connexes disjointes subira l'effet d'un déni de service. Puisque la métrique utilisée en survivabilité est la performance tel que perçue par l'usager, ce type de panne est beaucoup plus grave qu'une panne simple.

La gravité d'une panne est directement reliée à un autre attribut : la masquabilité d'une panne. La masquabilité est la possibilité de rendre une panne transparente à l'usager. Si une panne survient sans que l'usager ne puisse la percevoir, cette panne est dite masquable. Ce masquage peut être effectué de deux façons : soit en cachant la panne, soit en convertissant la panne en une panne jugée moins sévère. Par exemple, un message HTTP erroné sera détecté et rejeté par le code de détection d'erreur (convertissant une erreur de type byzantin en une erreur de type omission) et l'application demandera une retransmission pour assurer la continuité (masquant complètement la faute.) Toutefois, le partitionnement et les pannes excédant la capacité de redondance du système ne peuvent être masqués.

2.1.3 Modélisation mathématique de la survivabilité

Pour diminuer le temps de réparation d'un réseau, l'automatisation des actions de restauration a été implantée. Il est possible depuis fort longtemps de concevoir un réseau

capable de corriger automatiquement toute panne simple. Un réseau doté d'une telle capacité est dit auto-regénérant (self-healing). Pour s'assurer qu'un réseau puisse se réparer automatiquement, on doit planifier le réseau de sorte que tous les nœuds aient un degré plus grand ou égal à 2 et que chaque chemin soit couvert par un chemin redondant. Finalement, il faut que le réseau dispose des mécanismes logiciels pour restaurer les chemins en réorganisant la topologie du réseau. De façon générale, pratiquement tous les réseaux distribués auto-organisants (distributed self-organizing networks) sont des réseaux auto-regénérants. Cette propriété représente l'objectif à atteindre en matière de survivabilité.

Pour atteindre cet objectif, nous disposons de deux possibilités : l'utilisation de la redondance ou l'utilisation de la diversité. L'utilisation de la redondance se rattache à tous les mécanismes de protection utilisant une pré-allocation des ressources. Le problème mathématique relié aux méthodes de pré-allocation consiste à trouver deux chemins disjoints de nœuds respectant les contraintes de la qualité de service entre la source et la destination. Ce problème est résolu avant l'occurrence de la panne et dicte les actions correctives à prendre lorsqu'une panne survient. Dans le cas de l'utilisation de la diversité, on utilise plutôt une réallocation des ressources restantes après la panne pour en limiter l'effet. Ce problème se nomme le problème d'allocation de capacité résiduelle et est résolu après l'occurrence de la panne.

Dans le cas d'une méthode préallouée, on doit allouer les ressources dès l'établissement de la communication. Ceci cause un gaspillage des ressources réseau s'il n'y a pas occurrence de panne. De plus, puisque l'allocation des ressources se base sur l'état du réseau au moment de l'établissement des communications, les méthodes préallouées sont peu dynamiques et réagissent mal aux changements brusques de topologie. Toutefois, puisque le problème de l'allocation de ressource a déjà été résolu, les actions nécessaires pour restaurer la communication sont moins complexes. Elles consomment donc moins de ressources logicielles et donnent une vitesse de restauration beaucoup plus rapide.

Dans le cas des méthodes d'allocation de ressources résiduelles, on fait l'allocation des ressources qui sont disponibles après l'occurrence de la panne. Ceci permet l'utilisation optimale des ressources réseau disponibles après la panne. Aussi, puisque les méthodes d'allocation de ressources résiduelles se basent sur l'état du réseau au moment précis de l'occurrence de la panne, elles sont très dynamiques et réagissent bien aux changements de topologies. Toutefois, le problème d'allocation de ressource doit être résolu après l'occurrence de la panne, ce qui augmente la complexité des actions à entreprendre. Cette plus grande complexité se traduit par une augmentation du temps de restauration et de la consommation des ressources logicielles du réseau.

2.1.4 Coordination des mécanismes de survivabilité

Les différences entre les mécanismes de survivabilité peuvent causer des problèmes de coordination. Ces problèmes surviennent particulièrement au niveau de la coordination verticale des mécanismes dans le modèle OSI. En effet, une panne dans le réseau peut déclencher une réponse à plusieurs niveaux simultanément. Selon les caractéristiques de ces réponses, il est possible que plusieurs mécanismes déclenchés en parallèle entrent en conflit. Un exemple de ce type de conflit est un réseau MPLS sur SONET qui implémentent chacun des mécanismes de protection par commutation du trafic vers un chemin de secours. Advenant l'occurrence d'une panne, il serait fâcheux que les deux mécanismes se déclenchent en même temps. Dans l'exemple mentionné, quelques millisecondes après que SONET ait commuté le trafic sur la fibre de protection, MPLS redirigerait le trafic vers un LSP de secours annulant ainsi les actions prises par le niveau SONET. Pour contourner ce problème, on utilise des minuteries qui déclenchent les actions de correction une fois que le temps de réparation des niveaux inférieurs est écoulé et que la panne subsiste.

2.2 Survivabilité de bas niveau

Il est possible d'implémenter la survivabilité à tous les niveaux du modèle OSI. Parmi les méthodes de survivabilité pouvant être implémentés au niveau de la couche

physique du modèle OSI, on retrouve les topologies survivables - comme les anneaux, les réseaux maillés et les méthodes hybrides - et les méthodes basées sur SONET.

2.2.1 Topologie en anneau

La topologie en anneau était la topologie privilégiée des planificateurs voulant assurer une fiabilité de 99,99%. Cet engouement s'explique premièrement par la faible complexité d'une topologie en anneau. En effet, l'organisation d'un topologie en anneau est simple : chaque nœud est connecté à exactement 2 voisins pour former un cycle passant exactement une fois par chaque nœud et chaque lien. La Figure 2.5 illustre une topologie en anneau avec 5 routeurs.

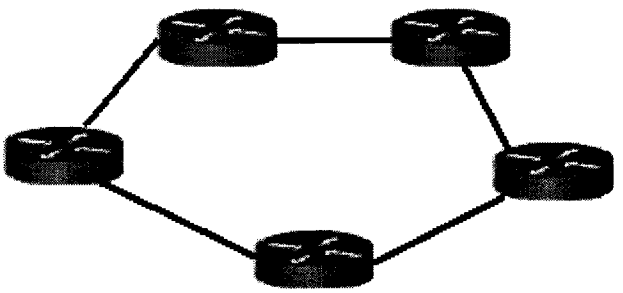

Figure 2.5 Topologie en anneau

Pour acheminer du trafic, il suffit de diriger le message sur le port rattaché au chemin le plus court. Pour assurer une redondance, il suffit d'envoyer une copie du message par le chemin le plus long. Cette faible complexité permet de limiter le coût en ressources logicielles puisqu'il est possible d'utiliser des nœuds ADM (add-drop multiplexers) plutôt que d'utiliser des routeurs. Un ADM se distingue d'un routeur par sa capacité limitée de traitement. Un ADM ne peut qu'ajouter, retirer ou faire suivre le trafic. Cette capacité limitée se traduit bien sûr par un coût de déploiement plus faible, mais un coût d'opération plus élevé.

L'opération d'une topologie en anneau est très peu efficace au niveau de la consommation de la bande passante. Cette inefficacité est largement causée par le trafic en transit. Grover [13] montre que des topologies en anneau, même si l'efficacité a été

optimisée dans le design, peuvent avoir besoin de jusqu'à trois fois la capacité d'une topologie utilisant le chemin le plus court pour la quantité de demande desservie. Dans le cas où l'anneau doit aussi assurer la survivabilité, l'efficacité est encore diminuée. En effet, comme l'indique [13], les topologies en anneau ont structurellement une redondance de 100% puisque, pour chaque unité de capacité opérationnelle, on retrouve une capacité de protection égale de l'autre côté de l'anneau.

Toutefois, cette inefficacité au niveau de la consommation des ressources se traduit par une grande rapidité de protection. La redondance de 100% assure que la protection est toujours disponible pour les cas de panne simple. Aussi, la résolution du problème d'affectation de capacité résiduelle est résolue avant la panne puisque le chemin de protection est déjà assigné. De plus, la complexité des opérations se réduit à inverser le sens de parcours du trafic. Les topologies en anneau sont donc capables d'enclencher la protection en un temps très rapide, généralement inférieur à 60 ms.

Une autre conséquence de la faible complexité est la faible adaptabilité des topologies en anneau par rapport aux types de fautes. Vu la topologie, il est évident que, dès que deux pannes surviennent en même temps sur des nœuds non adjacents, on se retrouve avec un partitionnement de réseau. De plus, ce type de panne a aussi pour conséquence d'annuler le chemin de protection pré-caclulé sans possibilité d'en calculer un nouveau. Cette faible complexité empêche aussi d'avoir la granularité nécessaire pour envisager une protection basée sur le type de trafic.

2.2.2 Topologie en maille

À l'opposé des topologies en anneau se situent les topologies maillées. Dans un réseau maillé, les nœuds sont connectés avec le plus grand nombre possible de voisins. Le degré de chaque nœud qualifie le réseau maillé. Dans le cas extrême, chacun des nœuds est relié à tous les autres nœuds du réseau pour former un réseau complètement maillé. La Figure 2.6 illustre un réseau complètement maillé avec 5 routeurs.

Les réseaux maillés sont plus complexes à gérer que les réseaux en anneau. Le coût en ressources logicielles pour acheminer le trafic (calcul des plus courts chemins)

est beaucoup plus élevé que dans le cas des topologies en anneau. De plus, ce coût est proportionnel à $n(n\text{-}1)$ où n est le nombre de nœuds. Ceci peut causer des problèmes d'évolutivité quand n devient grand. Finalement, le déploiement des liens entre tous les nœuds peut s'avérer coûteux au niveau aussi bien du coût des fibres que des droits de passage dans des zones urbaines.

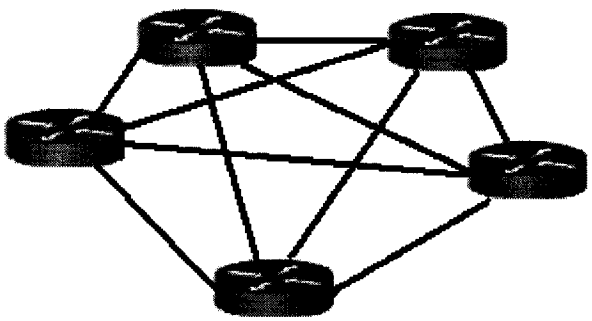

Figure 2.6 Réseau complètement maillé

Cette complexité permet toutefois d'utiliser avec un maximum d'efficacité la bande passante disponible. En effet, dans un réseau complètement maillé, il n'y a pas de problème de réacheminement de trafic puisque tous les nœuds sont adjacents entre eux. De plus, dans le déploiement d'un réseau maillé survivable, il est possible d'utiliser la capacité résiduelle sur chacun des liens pour restaurer le trafic. Ainsi, en utilisant la diversité inhérente des réseaux maillés, il est possible d'assurer une survivabilité avec beaucoup moins de ressources matérielles. Les réseaux maillés peuvent atteindre une survivabilité complète avec une redondance inversement proportionnelle au degré moyen des nœuds dans le réseau [18]. Ainsi, avec un degré moyen entre 3 et 4.5, il est possible d'assurer une survivabilité avec aussi peu que 30 à 50% de redondance.

Toutefois, cette consommation efficace de la bande passante a un prix. La complexité des réseaux maillés nécessite des calculs beaucoup plus longs pour réacheminer le trafic vers des liens fonctionnels. De plus, puisque les réseaux maillés utilisent la capacité résiduelle, quantité qui varie dynamiquement, il est impossible de préallouer la capacité résiduelle. Ainsi, à chaque occurrence de panne, on doit ajouter le

temps de résolution du problème d'allocation de capacité résiduelle au temps de restauration. Les techniques de restauration dans un réseau maillé sont inévitablement plus lentes que celles dans les réseaux en anneau.

2.2.3 Topologies hybrides

Dans le but d'obtenir un réseau avec les avantages des deux topologies, plusieurs propositions ont été soumises pour des solutions hybrides de maille et d'anneau. Certaines solutions proposent de diminuer la redondance inhérente d'une topologie en anneau pour augmenter l'efficacité au niveau de la consommation des ressources, tout en conservant la rapidité de l'anneau. D'autres solutions proposent des méthodes pour allouer la capacité résiduelle en avance dans les réseaux maillés afin d'améliorer la rapidité de protection sans trop dégrader la consommation de ressource. La topologie d'anneau augmenté (enhanced ring) est un exemple de solution du premier type et les cycles-p (p-Cycles) sont un exemple de solution du deuxième type.

L'idée générale de la topologie en anneau augmenté est de permettre à deux anneaux possédant un lien commun de protection de partager ce lien à des fins de protection. La Figure 2.7 illustre une topologie de protection dans une topologie de type anneau augmenté.

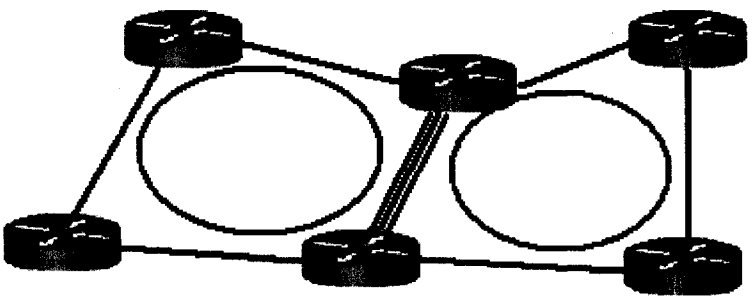

Figure 2.7 Topologie en anneau augmenté

L'utilisation du lien commun de protection permet de diminuer la redondance du réseau puisqu'on a un seul lien plutôt que 2 liens de protection en parallèle. Toutefois, cette méthode se base fortement sur l'hypothèse de l'indépendance des pannes.

La topologie en cycle-p se base sur l'allocation de cycles préétablis pour la protection mais permet au trafic d'être routé par le plus court chemin. On se retrouve alors avec un réseau maillé possédant un chemin de secours en anneau pré-alloué. La Figure 2.8 illustre un exemple de topologie en cycle-p :

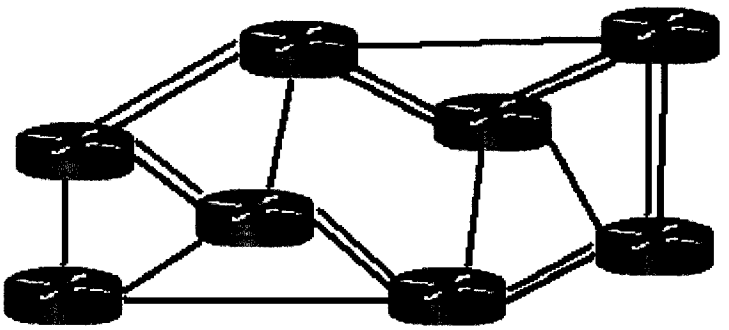

Figure 2.8 Topologie en cycle-p

Puisque les cycles sont formés en avance, on peut économiser le temps de résolution du problème d'allocation de la capacité résiduelle et atteindre une vitesse de protection équivalente à celle d'une topologie en anneau. Toutefois, cette méthode utilise une technique de protection de type anneau, ce qui peut causer des problèmes en l'occurrence de pannes multiples.

De nombreux autres modèles de topologie fiable ont été présentés dans la littérature (par exemple [27] et [32]). Malheureusement, appliquer ces méthodes aux pannes catastrophiques produirait des solutions très coûteuses. Cet axe de recherche n'a donc pas été privilégié.

2.2.4 Survivabilité SONET

En plus des avantages proposés par les topologies, l'architecture SONET possède ses propres mécanismes pour assurer la survivabilité. En effet, le protocole APS (Automatic Protection Switching) permet de modifier (soit automatiquement, soit manuellement) les paramètres des fibres de protection pour restaurer le signal en cas de panne. Les commandes ASP (automatic switch to protection) et ASW (automatic switch

to working) permettent de changer l'état d'une fibre si une perte de signal (ou le retour du signal) est détecté[2]. Cette méthode permet donc de détecter rapidement les bris, ce qui accélère d'autant plus la rapidité de protection. Toutefois, puisque cette commutation repose sur la perte du signal optique, il est impossible de détecter les pannes de nœuds. De plus, la protection se situe au niveau de la fibre. Ainsi, il est impossible d'implémenter une protection intégrant les classes de trafic car la granularité est insuffisante. Ces avantages et inconvénients sont typiques de mécanismes de protection se situant au niveau 1 du modèle OSI.

En plus de se situer au niveau 1 du modèle OSI, les mécanismes de protection SONET possèdent d'autres caractéristiques. En étant capable de changer de façon automatisée le statut d'une fibre de « protection » à « working » et vice-versa, les réseaux SONET ont la capacité d'être dynamiquement reconfigurables. Il est donc possible de faire un design de réseau auto-regénérant. Des exemples typiques de réseaux auto-regénérants basés sur l'architecture SONET sont les anneaux UPSR et BLSR. Ces deux architectures ajoutent les caractéristiques de mécanismes de protection en anneau aux avantages et inconvénients de SONET. En fait, de façon générale, les mécanismes de protection de l'architecture SONET reposent sur la protection par commutation vers une fibre de secours. Ces mécanismes possèdent ainsi les caractéristiques propres aux méthodes avec des ressources pré-allouées. Citons la rapidité de la protection, la mauvaise utilisation des ressources réseau, la faible complexité des opérations à effectuer et une mauvaise adaptation face aux pannes multiples.

Il nous a été possible de remarquer les caractéristiques des méthodes de protection pré-allouées et dynamiques ainsi que des particularités des mécanismes basé sur le niveau 1 du modèle OSI. Ces informations permettront de mieux évaluer les mécanismes de protection basés sur les couches intermédiaires du modèle OSI.

---

[2] Voir Annexe 1

2.3 Survivabilité de niveau intermédiaire

Les mécanismes de survivabilité des protocoles de couches 2 et 3 se distinguent des mécanismes de niveau 1. Nous étudierons donc les mécanismes de survivabilité basés sur le niveau IP pour illustrer les particularités des mécanismes de survivabilité de niveau 3 et sur MPLS pour illustrer les mécanismes basés sur le niveau 2.

2.3.1 Survivabilité IP

La survivabilité au niveau de la couche IP est principalement basée sur la conservation de la connectivité et la distribution de l'intelligence. Cette distribution permet au réseau de pouvoir résoudre le problème de réallocation des ressources (dans ce cas-ci, l'allocation des routes) malgré la perte de nœuds. Cette réallocation se fait par la résolution de problèmes de plus courts chemins basés sur l'information fournie par le réseau aux protocoles de routages L'éventuelle convergence de ces protocoles permet à un réseau IP de résister à n'importe quelle attaque ne brisant pas sa connectivité[3]. Toutefois, ces protocoles de routage ont un temps de convergence très long, de l'ordre des minutes. Cette échelle de temps n'est évidemment pas adaptée aux contraintes de qualité de service des applications évoluées. Pour contourner cette lacune, on utilise généralement un protocole de couche 2 capable d'implémenter la qualité de service comme MPLS ou ATM.

2.3.2 Survivabilité MPLS

La couche MPLS, GMPλS et la couche ATM sont parfois appelées la couche chemin (path layer). Cet abus de langage est motivé par la fonction de ces deux protocoles. ATM et MPLS possèdent tous deux la propriété de contrôler les chemins de bout en bout. Ce contrôle s'effectue par le biais des LSP sur MPLS et des VC sur ATM. Il est donc possible d'implanter la survivabilité avec une granularité de chemin à l'aide de ces protocoles. MPLS a été choisi pour illustrer les caractéristiques des mécanismes

---

[3] Voir annexe 2

de survivabilité basés sur le chemin. Cette section ne présente qu'un aperçu de la survivabilité par MPLS, le lecteur est encouragé à consulter [9], [11], [17], [19], [20], [23] et [24].

Le premier intérêt d'une architecture de survivabilité au niveau chemin est la possibilité d'ajuster les mécanismes de protection en fonction de la classe de trafic. Par exemple, sur MPLS, il est possible de configurer les LSPs avec un mécanisme de protection différent. Il est donc possible de planifier les actions de restauration en fonction des mêmes critères que pour l'établissement de FECs (c'est-à-dire les différents requis de qualité de service, le prix que paient les différents utilisateurs, la location de la source ou de la destination, etc.). On peut donc utiliser des actions de protection différentes selon la classe de trafic. Ceci permet de faire une meilleure utilisation des ressources réseau ayant survécu à la panne, puisqu'il sera possible d'accorder les ressources en fonction de la capacité des requis de qualité de service.

Cette utilisation plus efficace des ressources réseau a un prix. Premièrement, la complexité des actions de restauration nécessite une plus grande consommation de ressources logicielles qui se traduit généralement par une plus grande signalisation. Ceci cause inévitablement une augmentation du délai par rapport à des mesures de protection de la couche physique comme les mécanismes de protection basés sur SONET. Aussi, le distancement de la couche physique rend la détection d'erreur plus difficile. Pour détecter une panne, il faut se fier à des messages de découverte appelés communément « *hello* messages ». Une panne est détectée lorsqu'un voisin soudainement cesse de répondre à ces messages. La rapidité de détection d'erreur est donc directement liée à la fréquence où les nœuds s'échangent les messages. Une plus grande fréquence améliorera la vitesse de détection, mais surchargera le réseau de messages de signalisation. Ce temps de détection s'ajoute au temps de restauration.

Les propriétés de la stratégie de résolution du problème d'affection de la capacité de protection s'ajoutent aux caractéristiques de mécanismes basés au niveau chemin. Dans le cas de MPLS, on retrouve la protection par commutation comme méthode de pré-allocation et le reroutage rapide comme méthode d'affectation de capacité résiduelle.

2.3.3 Protection par commutation

La protection par commutation est une méthode de protection globale, c'est-à-dire que le chemin est protégé dans son entièreté par un chemin de protection. Lors de l'établissement d'un LSP, le routeur de tête (Ingress) signale le chemin actif ainsi qu'un chemin de protection disjoint du chemin actif. Lorsqu'une faute survient sur le chemin actif, un message d'indication de faute, ou FIS, est envoyé jusqu'au routeur de tête. Sur réception de ce message, le routeur de tête commute le trafic du chemin actif vers le chemin de secours. Il est aussi possible de faire remonter le trafic couramment en transit vers l'ingress pour éviter la perte de trafic. La Figure 2.9 [14] illustre une protection par commutation. On remarque le chemin actif en rouge et le chemin de protection en vert.

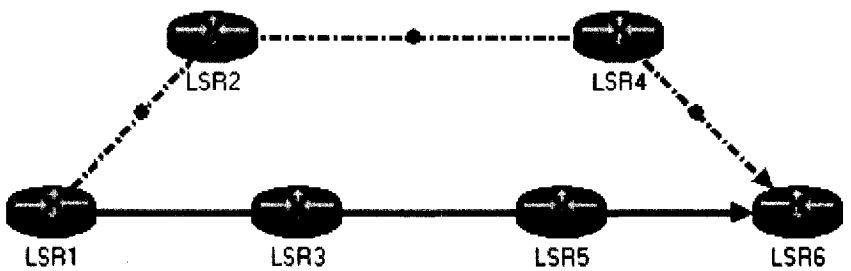

Figure 2.9 Protection par commutation

Comme les topologies en anneau, ce type de protection est intrinsèquement redondant à 100%, consommant ainsi beaucoup de ressources. Pour limiter cette consommation, plusieurs modes de fonctionnement ont été établis. Les modes principaux sont :

- 1+1 : Dans ce mode, le trafic est envoyé simultanément sur le chemin actif et le chemin de protection. Le nœud d'arrivée ne considère que l'information reçue par le chemin actif, sauf en cas de panne où le nœud d'arrivée ne considère que l'information provenant du chemin de secours.

On atteint ainsi la plus grande vitesse de protection, mais en consommant la plus grande quantité de ressources réseau.

- 1:1 : Dans ce mode, le trafic est envoyé sur le chemin actif et est commuté sur le chemin de secours advenant une panne. Il est donc possible de faire transiter du trafic de priorité moindre sur le chemin de secours. Ce trafic pourra être mis de côté si la commutation de protection doit être activée.
- n:m : Dans ce mode, n LSPs actifs partagent les mêmes m LSPs de protection. On peut alors diminuer la redondance au prix de limiter la quantité de pannes simultanées qu'il est possible de restaurer.

Aussi, similairement aux topologies en anneau, puisque le chemin de secours est calculé à l'avance, ce type de protection est très peu dynamique et s'adapte mal aux changements rapides de l'état du réseau.

2.3.4 Protection par reroutage rapide

Alors que la protection par commutation est une méthode basée sur la préallocation, le reroutage rapide est basé sur l'allocation dynamique de capacité résiduelle. Le mécanisme de protection est initié par le point de détection de la panne (PLR). Il n'est donc pas nécessaire de faire remonter le message de signalisation jusqu'à l'ingress, ce qui augmente la rapidité de la détection. Une fois la panne découvert, on signale un nouveau tunnel entre le point de réparation locale, ou PLR, et le point de réparation suivant la panne, ou PML. Ce tunnel de protection est agrégé au tunnel fonctionnel, ce qui a pour effet de rerouter le chemin emprunté par le trafic autour de la panne. La Figure 2.10 illustre un exemple de protection par reroutage rapide [14].

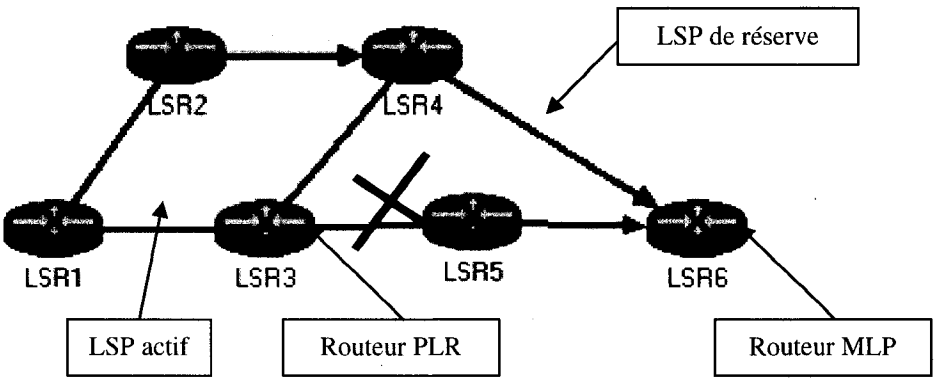

Figure 2.10 Protection par reroutage rapide

L'allocation dynamique des ressources permet de se baser sur l'état du réseau au moment de la panne. Il est donc possible d'optimiser la consommation des ressources réseau. Par contre, la méthode de reroutage rapide est plus longue que la méthode de protection par commutation. Aussi, pour assurer une protection, tous les routeurs doivent disposer de l'intelligence nécessaire pour effectuer les opérations de restauration.

Un problème propre au reroutage rapide est le problème de distribution des étiquettes. En mode de fonctionnement normal, il n'est pas nécessaire pour un routeur de connaître la stratégie de distribution d'étiquette des autres routeurs. Les seules informations essentielles, en mode de fonctionnement normal, sont les informations associant l'arrivée, sur un port P, d'un message marqué d'une étiquette X à la sortie, sur un port Q, du message avec une étiquette Y. Il est donc impossible d'utiliser ce mode de fonctionnement puisque l'étiquette de sortie du PLR n'est pas cohérente avec l'étiquette d'entrée du MLP. Pour contrer ce problème, l'utilisation du paramètre RRO (Record_Route_Object) du protocole RSVP-TE peut être utilisé. Cet objet contient une liste des LSRs et des étiquettes que ces LSRs utilisent pour le chemin associé au RRO. En utilisant comme étiquette de sortie de PLR l'étiquette d'entrée du MLP, il n'y aura pas de problème de cohérence. La Figure 2.11 illustre ce mécanisme [14].

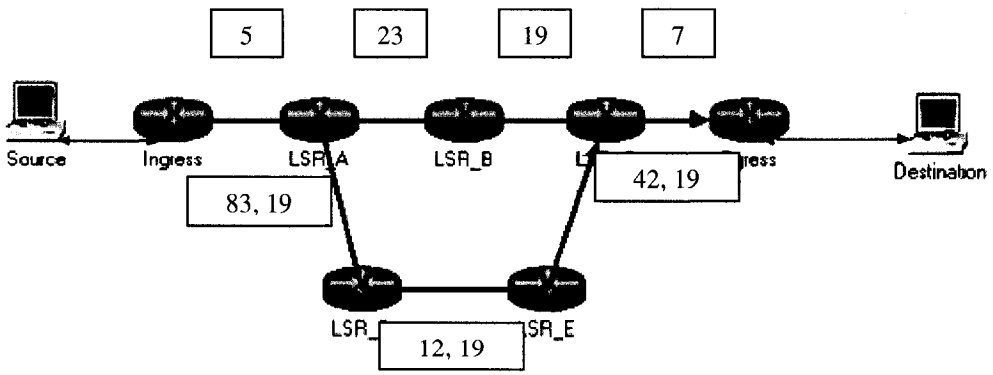

Figure 2.11 Problème de cohérence des étiquettes

Il est aussi possible d'éviter ce problème en redistribuant les étiquettes pour le chemin, mais cette solution s'avère beaucoup plus coûteuse au niveau de la signalisation et du temps d'exécution.

2.3.5 Modèle de RD-QoS

Le modèle de RD-QoS, ou « Resilience Differentiated Quality of Service », est conçu pour l'incorporation de la qualité de service aux actions de restauration. L'idée derrière l'architecture RD-QoS est d'associer une méthode de protection à une classe de qualité de service. Par exemple, un trafic de classe conversationnelle tolérant très mal le délai, est associé à une méthode de restauration utilisant la pré-allocation. La rapidité des techniques de restauration pré-allouées permet de rencontrer les exigences élevées de qualité de service de la classe conversationnelle. Pour le trafic de type meilleur-effort, des techniques utilisant l'allocation dynamique de capacité résiduelle seront utilisées pour optimiser l'utilisation des ressources et restaurer une plus grande quantité de trafic. Le coût additionnel en temps est négligé parce que la classe meilleur-effort ne possède pas de requis rigide de délai.

Tableau 2.1 Classes de résilience

Service class	RC1	RC2	RC3	RC4
Requis de résilience	Élevé	Moyen	Faible	Aucun
Temps de recouvrement	10 – 100 msec	100 msec – 1sec	1 sec – 10 sec	N.A.
Mécanismes de résilience	Protection	Restauration	Reroutage	Préemption
Mise en place du chemin de relève	Pré établi	Sur demande (Immédiatement)	Sur demande (Avec délai)	Aucun
Allocation des ressources	Réservé à l'avance	Sur demande (Avec garantie de disponibilité)	Sur demande (Si disponible)	Aucun
Qualité de service après le recouvrement	Equivalant	Peut être réduit temporairement	Peut être réduit	Aucun

Pour obtenir la correspondance, Autenrieth et Kirstädter (voir [2] à [8]) ont associé à chaque type de trafic une « classe de resilience ». Le Tableau 2.1 résume les caractéristiques de chacune des classes proposées par Autenrieth et Kirstädter ([2] à [8]).

L'article publié par Autenrieth et Kirstädter ([2] à [8]) fournit aussi les relations existant entre les classes de résilience et les classes de qualité de service ainsi que les extensions au protocole RSVP-TE nécessaires à l'implémentation de leur architecture.

2.4 Survivabilité de haut niveau

La survivabilité peut aussi être implémentée dans les couches supérieures du modèle OSI ou même en dehors de celui-ci. Les particularités des mécanismes de

survivabilité de haut niveau ainsi que trois exemples, les architectures réparties, les réseaux actifs et les applications reconfigurables, seront détaillés dans cette section.

2.4.1 Particularités de la survivabilité de haut niveau

La survivabilité de haut niveau est associée à des mécanismes évolués, basés sur des actions complexes. Cette complexité, conséquence directe de la plus grande disponibilité des ressources logicielles à mesure qu'on s'éloigne de la couche physique, est le principal point commun des mécanismes de survivabilité réseau de haut niveau. On peut donc discerner les conséquences de cette complexité dans tous les mécanismes de survivabilité entrant dans cette catégorie.

La principale conséquence de la plus grande complexité est la consommation accrue de ressources logicielles. Cette plus grande consommation est directement liée au plus grand niveau d'intelligence requis pour traiter cette complexité. Ce plus grand niveau d'intelligence nécessite une grande quantité d'information et de traitement, ce qui consomme de la mémoire dans le cas d'applications centralisées, ou augmentent la signalisation dans le cas d'applications réparties. L'augmentation du traitement (ou de la signalisation selon le cas) se traduit inévitablement par un temps de restauration plus lent. Ces désavantages représentent les limitations majeures des mécanismes de survivabilité de haut niveau.

Par contre, l'augmentation de l'intelligence permet d'affiner la granularité du traitement. Cette granularité plus fine se transpose aussi au niveau des actions de restauration qu'il est possible de poser. Par exemple, au plus haut niveau de complexité, il serait possible de poser des actions de survivabilité par message plutôt que par fibre, ou même classe de trafic. Ces actions plus fines permettent de faire un usage optimal des ressources réseau. En effet, puisque les actions de restauration permettent de rerouter des fractions de trafic et de poser des actions différentes selon l'information traitée, il n'est pas nécessaire d'avoir un degré de redondance élevé.

En ayant noté que les mécanismes de survivabilité de haut niveau sont plus complexes, consomment plus de ressources logicielles et temporelles, possèdent une

granularité plus fine et consomment moins de ressources réseau, on peut étudier les particularités des différents mécanismes.

2.4.2 Architectures réparties

Les architectures réparties consistent à diviser une architecture centralisée pour « distribuer » les ressources sur un réseau. Cette distribution se traduit généralement par une répartition géographique plus diverse. Ceci réduit grandement le type de fautes capables de causer des dommages au réseau. En effet, tous les composants d'une architecture centralisée fait partie du même GRP pour toutes les fautes qui dépendent de la géographie. Dans une architecture répartie, chacun des éléments est membre d'un GRP différent. Ainsi, le système complexe résultant de l'ensemble des éléments est beaucoup plus résistant aux pannes de ses composantes puisqu'une panne affectant une composante a moins de risque d'affecter les autres éléments du système complexe. Cette technique est appelée mitigation de risque. Nous verrons, à travers l'exemple des architectures multi-tiers et des bases de données réparties, comment s'effectue cette mitigation.

L'architecture multi-tiers est une amélioration de l'architecture client-serveur classique. L'idée est d'insérer entre le client et le serveur un tiers pour effectuer le contrôle de l'interaction avec l'usager. Ceci permet l'opération asynchrone de la part du client pendant que le serveur traite l'information. Cette opération asynchrone permet aussi de masquer les pannes au client pendant que le système est en train d'exécuter des actions de restauration. Ceci permet de mitiger l'influence des pannes du serveur.

Les bases de données réparties utilisent aussi une division pour mitiger l'influence des pannes. Par contre, cette distribution est plutôt géographique. En effet, une base de données répartie est une base de données qui a été répartie sur plusieurs nœuds du réseau. Chacune de ces partitions possède une partie (pouvant être la totalité) de l'information. Advenant l'occurrence d'une panne, on peut continuer à opérer avec l'information disponibles dans l'ensemble des bases de données qui ne sont pas affectées par la panne. Par exemple, dans le cas où on a quatre bases de données contenant

chacune 25% de l'information, il est possible de fonctionner après la panne d'une des bases de données avec 75% de l'information. Toutefois, ce mode de fonctionnement amène de nouveaux défis pour le développement de la base de données. Notamment, mentionnons la nécessité d'intégrer des mécanismes pour assurer l'intégrité et la cohérence des données. Le principal problème est de s'assurer de l'état des bases de données miroir au moment de la panne.

2.4.3 Réseaux actifs

Dans la même idée de mitigation des effets des pannes, des chercheurs se sont penchés sur la mitigation des pannes de l'administration réseau. Compte tenu de la nature centralisée des méthodes traditionnelles de gestion de réseau, plusieurs chercheurs favorisent une approche répartie à la gestion du réseau [11]. Une des méthodes pour gérer les réseaux de façon décentralisée est l'utilisation des réseaux actifs.

L'intérêt principal des réseaux à gestion décentralisée n'est pas directement lié à la survivabilité. Un réseau administré centralement est généralement incapable de gérer une grande quantité d'information sur le réseau. Ceci s'explique par la grande quantité de données inutiles (ne rapportant aucun changement significatif dans le réseau) transmis. Il serait possible de faire l'agrégation et de filtrer ces messages localement dans une architecture décentralisée. Ceci permettrait d'améliorer les actions prises à la suite de la découverte d'une panne et d'augmenter indirectement la survivabilité.

Les réseaux décentralisés permettent d'augmenter plus directement la survivabilité d'un réseau à des pannes multiples et corrélés en offrant la possibilité de prendre certaines décisions localement. Plusieurs décisions, basées sur l'état d'une petite partie du réseau, n'ont pas besoin d'être prises centralement. En prenant ces décisions localement, il est possible de réagir plus rapidement à des changements. Ceci permet à un réseau de s'adapter à des changements significatifs du réseau. Aussi, une répartition de la prise de décision permet de survivre à un partitionnement du réseau. Pour distribuer le management et la prise de décision, on considère l'utilisation de réseaux actifs.

Un réseau actif est un réseau programmable. Dans un paradigme de réseau actif, des programmes peuvent être injectés dans les équipements les rendant ainsi actifs dans le sens que la façon dont ces équipements fonctionnent et la façon dont ils traitent l'information peuvent être contrôlées et personnalisées dynamiquement [11]. Ainsi, les nœuds d'un réseau actif ne sont pas seulement des entités qui font suivre l'information, mais ils peuvent aussi manipuler l'information. En permettant ce traitement personnalisé et la possibilité de reprogrammer le réseau, les réseaux actifs constituent une infrastructure à nœuds ouverts. Le réseau peut ainsi supporter tout type de protocole et de services.

Une première approche pour implémenter ce type de réseau est l'approche discrète, aussi nommée « programmable switch approach ». Dans cette approche, les nœuds n'acceptent d'être reprogrammés que par des administrateurs. L'injection de nouveau programme ne peut être faite que par des canaux hors-bande. Ainsi, on sépare l'injection de nouveau code et la transmission des paquets. Il est donc possible d'assurer une meilleure authentification des usagers injectant du code et de vérifier le code qui a été injecté.

On peut trouver la procédure pour faire suivre un message en [29]. Tout d'abord, un administrateur programme les nœuds en injectant les routines nécessaires par le canal de contrôle. Un usager du réseau peut ensuite envoyer ses paquets sur le réseau comme avec un réseau traditionnel. Lorsque ce paquet atteint un nœud actif, le nœud analyse l'en-tête du paquet et détermine quel programme il doit appliquer à ce paquet. Ensuite, il lance les programmes appropriés sur le paquet et, une fois que toutes les opérations sont terminées, il fait suivre le paquet selon les instructions du programme d'acheminement.

La seconde approche pour l'implantation des réseaux actifs est l'approche intégrée. Dans le mode intégré, l'injection des programmes et le traitement des messages ne sont pas séparés. L'intelligence se retrouve ainsi au niveau des paquets et non des nœuds. Le réseau est donc beaucoup plus ouvert, mais beaucoup moins sécure puisqu'il est beaucoup plus difficile d'authentifier l'usager et de déterminer si les instructions sont dangereuses.

La procédure pour faire suivre les paquets dans un modèle intégré est décrite en [29]. Tout d'abord, le message est encapsulé par l'usager dans une en-tête incluant le code devant être appliqué sur le paquet et est envoyé sur le réseau. Une fois arrivé à un nœud actif, un processus capable d'identifier les extrémités du paquet est lancé pour extraire le paquet du flot binaire circulant sur le lien. Le paquet est ensuite placé dans un environnement temporaire de traitement pour évaluer de façon sécuritaire le code contenu dans le paquet et les permissions de l'usager d'où provient le paquet. Une fois que le paquet est déclaré inoffensif, le nœud effectue les opérations demandées dans le code, ce qui peut inclure de faire suivre le paquet.

Finalement, une troisième implémentation des réseaux actifs est possible. Dans cette implémentation, les administrateurs programment les nœuds pour y introduire certaines primitives. Ces primitives sont utilisées par les usagers envoyant des paquets contenant du code faisant appel à ces primitives. Cette implémentation est donc une hybridation des modes discrets et intégrés.

2.4.4 Applications reconfigurables

Les applications reconfigurables se distinguent des autres mécanismes de survivabilité de haut niveau par l'utilisation de techniques classiques de survivabilité plutôt que l'utilisation de techniques de mitigation. L'architecture RAPTOR introduite dans la thèse de doctorat de Matthew C. Elder [15] est un excellent exemple d'implémentation de mécanisme de survivabilité au niveau de la couche application.

L'architecture RAPTOR vise à protéger les réseaux critiques d'information contre des attaques de grande envergure. Dans ce type d'attaque, il n'y a pas suffisamment de ressources réseau pour maintenir un niveau de fonctionnement régulier. Pour protéger les services de type mission critique et qui ne peuvent pas tolérer d'interruption, il faut alors reconfigurer les services offerts par les applications. Cette reconfiguration s'effectue en entrant un mode alterné de service qui supporte une dégradation du service offert. La Figure 2.12 [15] illustre l'architecture utilisée pour reconfigurer les services offerts par le réseau.

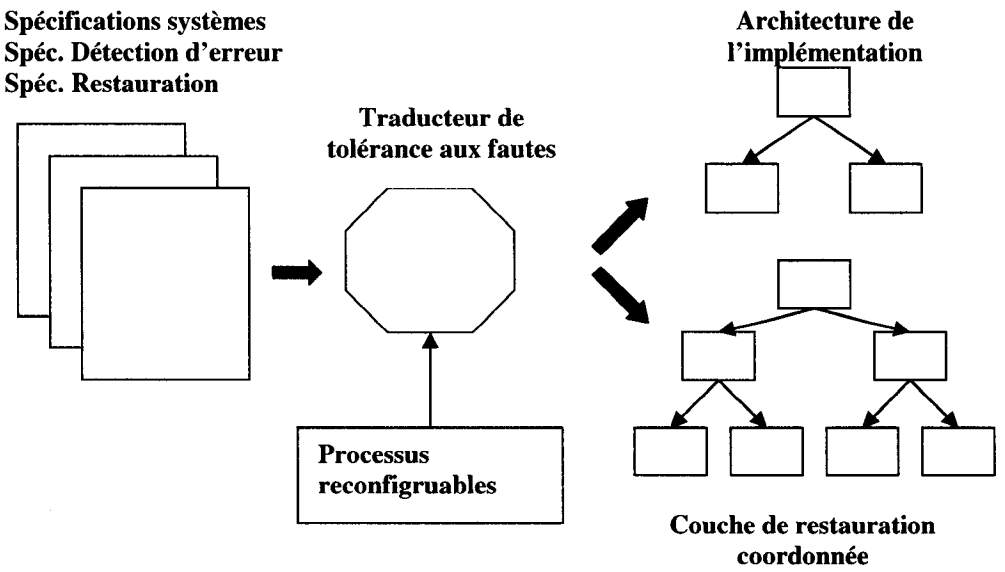

Figure 2.12 Architecture de restauration RAPTOR

Dans cette architecture, les différents éléments envoient leurs spécifications et les spécifications de l'erreur à un système de traduction. Ce système de traduction évalue ensuite l'information pour déterminer quelles actions doivent être prises et quels services sont critiques. Une fois que la décision est prise, elle est envoyée au système de contrôle qui va initier la restauration du réseau.

Pour reconfigurer une application, l'architecture RAPTOR se base sur des processus ayant la possibilité d'être stoppés à des endroits propices. Ces endroits doivent être assez fréquents pour permettre un temps de réponse adéquats, mais ils doivent aussi assurer la consistance de l'état. Ces processus doivent aussi posséder la possibilité de démarrer de nouvelles fonctionnalités, tout en restant consistants avec l'état du processus au moment de son arrêt. Ainsi, une action de restauration peut redémarrer un processus, stopper un processus, ou redémarrer un processus avec les fonctionnalités du « mode panne ». Les applications peuvent ainsi être reconfigurées pour offrir un service alternatif en cas de panne qui consommerait moins de ressources réseau. Les

applications ne pouvant souffrir d'interruption peuvent donc poursuivre leurs opérations et les applications non vitales pourront être reprises une fois que la panne sera réglée. La granularité de protection de l'architecture RAPTOR se situe donc au niveau du processus.

Cette granularité extrêmement fine du processus dicte sa performance, puisqu'une décision d'action corrective doit être prise pour chaque processus. Le temps de restauration est très lent. En plus d'être lent, il requiert beaucoup de ressources logicielles. Par contre, en ne restaurant que les processus vitaux, on atteint une consommation optimale des ressources réseau. Cette consommation optimale permet de maximiser la quantité de trafic qui est restaurée.

2.5 Conclusion

Dans ce chapitre, nous avons recensé et analysé les mécanismes de survivabilité réseau présent dans la littérature. Cette analyse nous a permis de faire la distinction entre les mécanismes utilisant une préallocation de ressources et les mécanismes utilisant une allocation dynamique de la capacité résiduelle. Dans le premier cas, les mécanismes sont rapides et peu complexes, mais aussi peu efficaces au niveau de la consommation de bande passante et peu résistants à des changements dynamiques. Dans le deuxième cas, les mécanismes font une meilleure utilisation de la bande passante et sont adaptés à des changements rapides de l'état du réseau. Par contre, ils sont moins rapides et plus complexes que les mécanismes basés sur la préallocation. Aussi, nous avons pu remarquer que la granularité du mécanisme de protection avait un effet direct sur le temps de restauration et la consommation de ressources réseau. Ces conclusions nous guideront pour le développement de mécanismes de survivabilité destinés à protéger les réseaux critiques d'information contre des pannes catastrophiques, ce que nous aborderons dans le prochain chapitre.

CHAPITRE III

MÉCANISMES DE SURVIVABILITÉ POUR FAILLE MAJEURE BASÉS SUR MPLS

De nos jours, la nécessité de protéger les infrastructures de communications est largement reconnue. Toutefois, la littérature ne propose aucune méthode permettant d'assurer la survivabilité des infrastructures de télécommunications face à des pannes majeures et corrélées. En effet, les méthodes utilisant la pré-allocation de ressources utilisent trop de ressources réseau dans les cas de pannes multiples et les méthodes de post-allocation ne sont pas assez rapides pour traiter du trafic avec des contraintes de qualité de service. Pour pallier ce problème, ce chapitre présente deux mécanismes de survivabilité pour les réseaux métropolitains basés sur MPLS et compare leurs performances. Les principes et fondements mathématiques de l'approche choisie seront présentés dans la première partie de ce chapitre. Nous poursuivrons en présentant une étude du délai et de la survivabilité des approches proposées ainsi que de la méthode de protection par commutation de MPLS. Ensuite, la modélisation de nos solutions, les détails de l'envergure du prototype et les paramètres d'implémentation seront finalement développés.

3.1 Principes et fondements de la solution proposée

Cette section présente les principes justifiant les choix de design ainsi que les fondements mathématiques de la solution proposée.

3.1.1 Principes

Il a été établi au chapitre 2 que les mécanismes de survivabilité existants ne sont pas adaptés au traitement des pannes majeures et corrélées dans des réseaux à trafic avec contraintes de qualité de service. Ces lacunes sont liées aux contraintes de délai,

auxhypothèses sur les pannes simples et à la consommation de bande passante. D'où les principales caractéristiques de la solution proposée, soient la différenciation de protection en fonction de la classe de trafic, la prédéfinition de chemin et la division physique de la bande passante de protection.

Le principal défi à relever est celui de la qualité de service. Étant donné le grand nombre de pannes, il est impossible d'empêcher la congestion dans le réseau. Pour éviter que cette congestion se traduise en délais inacceptables pour les classes possédant des contraintes strictes de qualité de service, il faut réserver des ressources. Cette réservation lie le problème de la congestion au problème de disponibilité des ressources. En effet, les ressources réservées afin de conserver la qualité de service limite la bande passante disponible, même si ces ressources ne sont pas utilisées. Cette consommation de bande passante est tout à fait acceptable en mode de fonctionnement normal. Toutefois, dans un scénario à ressources limitées comme un scénario de panne majeure, il faut limiter le gaspillage de bande passante au maximum afin de pouvoir maximiser la quantité de trafic restaurée.

Pour économiser la bande passante, on peut remarquer qu'il y a deux grands types de trafic : les trafics sensibles aux délais, et les trafics non sensibles aux délais. La grande partie des trafics sensibles aux délais dans les réseaux de prochaine génération sont des trafics de type conversationnel (voice-over-IP et vidéoconférence). Les requis de délais sur ces trafics sont très stricts (150 ms pour la voix), mais ces types de trafic tolèrent une certaine perte de paquets. Dans le cas du trafic non sensible aux délais, on retrouve principalement des données. Le trafic de données est à l'inverse du type conversationnel puisque les délais sont tolérés, mais la perte de paquet ne l'est pas. Il est donc plus important de préserver les données que de les acheminer rapidement.

Ainsi, pour assurer la qualité de service des classes de trafic sensibles aux délais sans utiliser trop de bande passante, la solution proposée s'inspire de [14] pour différencier le traitement de deux classes de protection : le trafic sensible au délai (C1) et le trafic non sensible au délai (C2). Le trafic C1 sera traité par un mécanisme de protection basé sur le mécanisme de protection par commutation de MPLS qui réservera

la bande passante de protection par RSVP-TE pour garantir une certaine qualité de service. Le trafic C2 sera traité par le mécanisme de protection de la couche 3, c'est-à-dire le reroutage IP basé sur OSPF. Le mécanisme du niveau 3 ne s'enclenchera qu'après la complétion du mécanisme de niveau MPLS et soustraira la bande passante réservée pour le trafic C1 de la largeur de bande du lien dans le calcul des métriques OSPF afin d'éviter que le trafic de type C2 ne se retrouve en famine dans les liens congestionnés par le trafic C1.

Pour C1, nous avons le choix entre deux méthodes : la pré-allocation et la post-allocation. Puisque le trafic C1 est plus sensible au délai, les méthodes de type pré-allocation sont attrayantes. Toutefois, tel que vu au chapitre 2, la pré-allocation est sensible aux pannes multiples. Pour des cas de pannes multiples, il faudrait réserver un grand nombre de chemins de secours. Ceci devient extrêmement coûteux en ressources réseau. Pour éviter le problème, nous aurons recours à la pré-définition de ressources. Similairement à la pré-allocation, les ressources de secours sont calculées et signalées avant la panne, mais, contrairement à la pré-allocation de ressources, les ressources ne sont pas réservées.

Le principal problème de la pré-définition de ressources demeure l'impossibilité de garantir la disponibilité des ressources au moment de la panne. Pour contrer ce problème, les solutions proposées suggèrent de faire une pré-allocation d'une partie des ressources. Dans la première solution, la bande passante b du chemin de protection est répartie sur n chemins. Ainsi, on se retrouve avec n chemins de protection (disjoints de nœuds) de bande passante b/n. On se retrouve donc avec la même bande passante garantie que dans un cas de pré-allocation de ressources, mais on dispose de n-1 chemins additionnels qui sont signalés advenant une panne sur le chemin de protection principal. Dans la deuxième solution, on réserve les n chemins de la même façon que dans la première solution. Lorsqu'il y a occurrence de panne, on envoie une demande d'élargissement de la bande passante sur les chemins pré-définis pour éviter de diviser le chemin et de réduire les performances.

Un autre avantage de la méthode proposée est la création d'un mode intermédiaire entre le meilleur cas et le pire cas. En effet, l'architecture de protection MPLS ne permet que le meilleur et le pire cas. Dans le meilleur cas, il n'y a pas de panne sur le chemin de protection et le temps de recouvrement est très faible. Dans le pire cas, il y a panne sur le chemin de protection en même temps que sur le chemin de travail, et le temps de recouvrement est très grand. Les solutions proposées permettent un certain nombre de cas entre le meilleur cas (aucune panne sur les chemins de protection) et le pire cas (tous les chemins de protection tombent en panne) où les performances seront partiellement dégradées (soit parce que la bande passante de protection n'est pas entièrement disponible, soit parce que le temps de recouvrement est plus long). Ce mode dégradé correspond à la conversion d'une panne impossible à masquer en panne de moindre conséquence.

En résumé, les solutions proposées se basent sur la pré-définition d'un grand nombre de chemins de secours alternatifs disjoints de nœuds et la différenciation entre deux grandes classes de trafic pour augmenter la survivabilité des réseaux à des pannes de grande envergure.

3.1.2 Fondements de l'approche

Dans cette section, nous évaluerons mathématiquement les deux solutions proposées pour la classe C1 en les comparant à la méthode de protection par commutation à partir de l'ingress par MPLS. Nous ferons ensuite une brève comparaison du mécanisme proposé pour la classe C2 et le mécanisme de restauration IP avec OSPF. Ces comparaisons s'attarderont au délai (en ms) comme métrique représentant la qualité de service et à la quantité de trafic restaurée (en Kbps) comme métrique représentant la survivabilité. Nous étudierons tout d'abord les hypothèses et conventions relatives à l'analyse mathématique. En deuxième lieu, nous comparerons une architecture de protection existante avec les architectures proposées. Premièrement avec la solution avec division de la bande passante (Solution1), ensuite avec la solution sans division de la bande passante (Solution2).

Hypothèses

Cette section détaille les hypothèses nécessaires à l'analyse mathématique ainsi que les notations utilisées pour l'évaluation des modèles.

Comme première hypothèse, seules les pannes de nœuds seront considérées. Ceci correspond au cas le plus commun dans une situation de panne catastrophique. Citons l'exemple d'un tremblement de terre qui provoque la destruction de bâtiments dans lesquels sont abrités les nœuds, mais ne cause pas de dégâts aux liens qui sont assez flexibles pour résister. Aussi, puisque la panne d'un nœud entraîne immédiatement la panne des liens qui lui sont attachés, le scénario de panne de nœuds est plus fort que le scénario de panne de liens.

Deuxièmement, une des hypothèses de base des études de tolérance aux fautes est l'hypothèse d'indépendance des pannes. Cette hypothèse est nécessaire à l'étude puisqu'elle permet d'évaluer la probabilité d'événements composés comme la probabilité de panne sur un chemin donné. Dans le cas des pannes catastrophiques, il est toutefois impossible de poser une telle hypothèse puisqu'il est connu que les pannes sont corrélées. De plus, il est pratiquement impossible d'évaluer la corrélation entre les pannes pour isoler les probabilités de panne sur un seul élément de réseau. Afin de faciliter l'étude, nous allons poser l'hypothèse suivante :

- La probabilité que tous les nœuds d'un réseau métropolitain sujet à un désastre tombent en panne est de 1, la corrélation entre les pannes est de 1.

Cette hypothèse est réaliste compte tenu qu'une abondance de pannes dans un réseau peut causer la panne des autres éléments en cascade jusqu'à ce que tous les nœuds tombent en panne. Toutefois, une certaine proportion de ces pannes pourra être recouvrée au niveau matériel. Ce recouvrement peut se faire par la mise en service automatique d'un élément secondaire mis en veille (backup) ou encore par le redémarrage des services ayant souffert d'une panne en cascade si l'élément a physiquement survécu à la catastrophe. On pose donc l'hypothèse suivante :

- Le temps pour réparer une panne recouvrable est négligeable (≈ 0).

Ce qui est réaliste puisque ce recouvrement s'effectue au niveau matériel, il s'agit d'un recouvrement extrêmement rapide. Pour les pannes qui ne peuvent pas être réparées instantanément, il faudra attendre un apport humain pour restaurer le service. On pose donc l'hypothèse suivante :

- Le temps pour réparer une panne non recouvrable tend vers l'infini ($\rightarrow\infty$).

Ce qui est réaliste puisque l'apport humain lors de circonstances catastrophiques peut prendre plusieurs jours, voire plusieurs semaines pour réparer. Comparativement à l'échelle de réparation automatique des éléments de réseau (qui est de l'ordre des minutes dans le pire cas), le temps de restauration est infiniment long. Puisque la survie des divers éléments de secours est un événement fortuit, il est possible de supposer que la probabilité de subir une panne recouvrable ne dépend pas de la probabilité de survie des autre éléments de réseau. Ainsi, on peut poser l'hypothèse suivante :

- La probabilité qu'un nœud subisse une panne recouvrable est indépendante de la probabilité que les autres nœuds du réseau subissent une panne recouvrable

Donc, bien que les probabilités de panne soient corrélées, les probabilités de survie des nœuds ne sont pas corrélées. Ceci nous permet de supposer l'indépendance de la survie des nœuds. Aussi, il nous est possible de supposer l'indépendance de l'événement composé qui est représenté par la survie d'un chemin donné. Ces hypothèses nous permettront de procéder à une étude analytique des architectures de survivabilité proposées.

Pour l'évaluation des divers modèles, les paramètres suivants seront utilisés :

- t_{PCC} : temps de calcul du plus court chemin
- t_{RIP} : temps de convergence des tables de routage
- t_{LSP} : temps de signalisation d'un LSP
- t_R : temps de restauration
- d : délai de transmission sur le chemin restauré
- s_i : temps de signalisation du LSP i

- t_T : temps de transmission
- t_{voix} : temps nécessaire pour restaurer tout le trafic de voix
- t_{TTR} : temps requis pour dévier le trafic sur le chemin de protection
- p_i : probabilité de survie du chemin de secours i
- p_s : probabilité de survie d'un nœud
- m : nombre de canaux de secours
- b_i : largeur de bande du LSP i
- q : quantité de trafic restaurée
- $|n|_{LSPi}$: nombre de nœuds sur le chemin i

3.2 Étude du délai

Dans cette section, nous analysons le délai des diverses méthodes de protection pour les deux classes de services principales.

3.2.1 Méthode de protection par commutation, classe C1

La méthode de protection par commutation à partir du nœud source constituera la référence pour comparer les architectures proposées. Nous évaluerons l'espérance mathématique du délai et l'espérance mathématique de la survivabilité associées à l'utilisation de ce mécanisme de protection.

Dans une architecture de protection avec pré-allocation (ou pré-définition), le délai d'un mécanisme de protection est composé de quatre délais :

1. Délai de détection de la panne
2. Délai de notification de panne
3. Délai de recouvrement
4. Délai de transmission sur le chemin restauré.

Les délais de détection et de notification sont identiques dans toutes les solutions évaluées. Ces délais ne seront donc pas pris en compte dans l'analyse. L'analyse

évaluera aussi le temps d'établissement du LSP. Bien que ce temps ne fasse pas partie de la phase critique (l'établissement du LSP précède la panne), l'étude de l'augmentation du temps d'établissement permet de prouver que les solutions proposées ne dégradent pas la qualité de service en mode de fonctionnement normal.

Le temps de signalisation d'un LSP est constitué de deux éléments. Tout d'abord, il faut calculer (ou regarder dans une table de routage) le chemin le plus court entre la source et la destination. Ensuite, il faut signaler (par RSVP) la réservation des ressources. On obtient alors l'équation (3.1) pour l'espérance du temps de signalisation d'un LSP pour la méthode de protection par commutation à l'ingress :

$$E(t_{LSP}) = t_{PCC} + s_i \qquad (3.1)$$

Pour l'espérance du temps de restauration, il faut prendre en compte deux possibilités. Dans la première possibilité, il n'y a pas de panne sur le chemin de secours lorsqu'il y a panne sur le chemin de travail. Il est donc possible de restaurer le trafic en déviant celui-ci sur le chemin de protection. Dans le deuxième cas, le cas où il y a panne à la fois sur le chemin de protection et le chemin de travail, la connexion sera perdue. Le délai sera donc de l'ordre du rétablissement de connexion. Comme ce délai est de l'ordre de minutes (voire des jours si le medium physique est détruit), il sera considéré comme infini. Ceci nous donne l'équation suivante :

$$E(t_R) = p_i \cdot t_{TTR} + (1 - p_i) \cdot \infty \qquad (3.2)$$

On remarque que l'espérance de ce délai est toujours infinie à cause de la présence de l'élément absorbant. Prenons donc note que dans le meilleur cas, le délai est de t_{TTR} ms et que dans le pire cas, le délai tend vers l'infini.

Finalement, le délai de transmission sur le chemin restauré ne dépend que de la topologie du réseau puisque la bande passante est réservée. Compte tenu de la notation utilisée, l'espérance du délai de transmission sur le chemin restauré sera donc :

$$E(d) = t_T \qquad (3.3)$$

Ce qui correspond à une transmission normale à travers un réseau.

Ces valeurs seront utilisées pour évaluer la performance des architectures de protection proposées.

3.2.2 Méthode avec division de la bande passante, classe C1

La méthode avec division de la bande passante utilise tous les canaux ayant survécu à la panne comme canal de protection. Cette configuration équivaut à diviser une même file d'attente par un nombre de serveurs égal au nombre de canaux de protection. Cette configuration garantit donc l'existence d'un service (en autant qu'au moins un canal de secours ait survécu), mais assure aussi une dégradation de la qualité de service.

Le temps de signalisation des deux architectures proposées dépend de l'algorithme utilisé pour calculer les m plus courts chemins entre la source et la destination. Par exemple, pour l'algorithme de calcul successif de m plus courts chemins disjoints, il faut compter m fois le temps de calcul d'un plus court chemin. Une heuristique considérant toutes les paires de chemins disjoints prendrait un temps différent. Pour simplifier les calculs, l'algorithme de calcul successif sera utilisé. Pour la signalisation par RSPV-TE, il n'est pas nécessaire de calculer n fois le temps de signalisation puisque ce processus peut se faire en parallèle. Il faudra plutôt considérer le temps maximal de la signalisation sur chacun des chemins. Nous obtenons alors l'équation suivante :

$$E(t_{LSP}) = m \cdot t_{PCC} + \max_{i \in [1,|n|]}(s_i) \qquad (3.4)$$

En comparant avec l'équation (3.1), on remarque que le temps d'établissement est plus long. Toutefois, la principale augmentation se situe au niveau du calcul des plus courts chemins, ce qui peut s'effectuer en un temps raisonnable. On peut donc conclure que la qualité de service n'est pas dégradée de façon significative dans un mode de fonctionnement normal.

Pour le temps de restauration, on doit considérer deux possibilités. Dans la première, il existe au moins un chemin de secours ayant survécu à la panne. Dans ce cas, le temps de restauration sera le temps d'effectuer la commutation sur les chemins de secours ayant survécu. Dans le second cas, il y aura coupure de la connexion, ce qui entraînera un temps de restauration tendant vers l'infini. On se retrouve donc avec l'équation suivante :

$$E(t_R) = (1-(1-p_1)(1-p_2)...(1-p_m)) \cdot t_{TTR} + (1-p_i)(1-p_2)...(1-p_m) \cdot \infty \quad (3.5)$$

À l'instar de la méthode par commutation standard, on se retrouve avec l'espérance du délai de restauration tendant vers l'infini. Aussi, le délai du meilleur cas est t_{TTR} ms et le délai du pire cas tend vers l'infini. Toutefois, en comparant avec l'équation (3.2), on remarque que le pire cas est $(1-p_2)...(1-p_m)$ fois moins probable dans la solution proposée. Ceci s'explique par le fait qu'un moins grand nombre de connexions seront bloquées.

L'évaluation analytique de l'espérance du délai de transmission dans l'architecture avec division de la bande passante est pratiquement impossible. En raison de la division de la bande passante, il est possible de perdre une certaine partie de la bande passante s'il y a panne sur un certain nombre de canaux de secours. Puisque le trafic ne disposera pas de toute la bande passante garantie, il est assuré que le délai sera dégradé. Dans un modèle M/M/1, la dégradation du délai dans la file d'attente n'est pas linéaire. Il est aussi impossible de faire des calculs précis de file d'attente puisque ceux-ci dépendent fortement de l'état du réseau. En effet, lorsque les liens utilisés sont peu chargés, le trafic marqué comme pouvant être échappé (c'est-à-dire le trafic dépassant la bande passante réservée) ne sera jamais rejeté. Ainsi, le délai souffrira peu. À l'opposé, dans un cas de congestion, il est possible que le délai tende vers l'infini. Le délai dans les files d'attentes dépend aussi des paramètres du trafic comme la quantité et la taille des rafales. Compte tenu de ces problèmes, il nous est possible de donner une valeur précise seulement dans le meilleur cas. En effet, le meilleur cas correspond à la division

du trafic sur m files d'attentes de capacité de service b_i/m. On obtient alors un délai m fois plus grand que lorsqu'une seule file d'attente est utilisée pour acheminer le trafic. On obtient donc l'équation suivante pour l'espérance du délai dans l'architecture avec division de bande passante :

$$E(d) = p_1 p_2 ... p_n \cdot m \cdot t_T + \textit{dégradation non linéaire} \tag{3.6}$$

Cette dégradation sera la principale source de dégradation du délai dans l'architecture avec division de bande passante.

3.2.3 Méthode sans division de la bande passante, classe C1

La méthode sans division de la bande passante utilise tous les canaux ayant survécu à la panne comme indication de chemin possible vers la destination. En cas de panne, une demande d'élargissement de la bande passante sera faite. Cette configuration permet d'éviter la division de la file d'attente qui s'accompagne inévitablement d'une multiplication du délai. Cette configuration ne garantit donc pas l'existence d'un service (si aucune bande passante additionnelle n'est disponible sur les canaux), mais ne dégrade pas la qualité de service.

Similairement à l'architecture avec division de la bande passante, le temps de signalisation dépend de l'algorithme utilisé pour calculer les m plus courts chemins entre la source et la destination. Toujours pour simplifier les calculs, l'algorithme de calcul successif sera utilisé. Identiquement, il est possible d'utiliser le parallélisme pour la signalisation par RSVP-TE. Nous obtenons alors l'équation suivante :

$$E(t_{LSP}) = m \cdot t_{PCC} + \max_{i \in [1,|n|]}(s_i) \tag{3.7}$$

En comparant avec l'équation (3.4), on remarque que le temps d'établissement est identique. Il est donc possible de tirer la même conclusion que pour la méthode avec division de la bande passante et affirmer que la qualité de service en mode de fonctionnement normal n'est pas dégradé significativement.

Pour le temps de restauration, on doit maintenant considérer plusieurs possibilités. Dans le meilleur cas, le premier canal de protection choisi pourra obtenir une augmentation de bande passante. Dans ce cas, le temps de restauration sera simplement le temps de commuter le trafic. Dans le pire cas, aucun LSP de protection ne survivra à la panne et il y aura perte de connexion. Ces situations correspondent aux deux possibilités existant dans les méthodes analysées plus haut. Dans l'architecture sans division de bande passante, il existe un certain nombre de cas intermédiaires. Ces cas correspondent à la commutation successive du trafic sur les chemins de secours restant, advenant une panne (bris d'équipement ou impossibilité d'obtenir de la bande passante additionnelle) sur les premiers chemins choisis. Dans ces cas, le temps de restauration équivaut aux temps de commutation successifs additionné aux temps d'attente pour déclencher la commutation. On se retrouve donc avec l'équation suivante :

$$\begin{aligned} E(t_R) &= p_1 \cdot t_{TTR} + (1-p_1)p_2 \cdot (2 \cdot t_{TTR} + t_T) + \ldots \\ &+ (1-p_1)\ldots(1-p_{i-1})p_i \cdot (i \cdot t_{TTR} + (i-1) \cdot t_T) + \ldots \\ &+ (1-p_i)(1-p_2)\ldots(1-p_m) \cdot \infty \end{aligned} \qquad (3.8)$$

Pareillement à la méthode par commutation standard, on se retrouve avec l'espérance du délai de restauration tendant vers l'infini. Aussi, le délai du meilleur cas est t_{TTR} ms et le délai du pire cas tend vers l'infini. Toutefois, en comparant avec l'équation (3.2), on remarque que le pire cas est $(1-p_2)\ldots(1-p_m)$ fois moins probable dans la solution proposée. Ceci s'explique par le fait qu'un moins grand nombre de connexion seront bloquées. En comparant avec l'équation (3.5), on remarque que le pire cas est également probable, mais que le meilleur cas est moins probable. Ceci indique que l'espérance du temps de restauration de la solution sans division de la bande passante est plus grande que celle avec division de la bande passante. Toutefois, la dégradation est linéaire et facilement quantifiable. Cette composante est la principale cause de la dégradation du délai dans l'architecture sans division de la bande passante.

Le délai de transmission dans l'architecture sans division de bande passante est formé de deux composantes. La première composante est la demande d'augmentation de bande passante. Cette composante est négligeable puisqu'elle est transmise attachée au paquet utile. La deuxième composante est le temps de transmission sur le chemin restauré. On obtient donc l'équation suivante pour l'espérance mathématique du délai dans l'architecture sans division de bande passante :

$$E(d) = t_T \qquad (3.9)$$

Tableau 3.1 Comparaison des délais pour le trafic de type C1

	Protection par commutation à l'ingress	Architecture avec division de la bande passante	Architecture sans division de la bande passante
t_{LSP}	$t_{PCC} + s_i$	$m \cdot t_{PCC} + \max_{i \in [1,\lfloor n \rfloor]}(s_i)$	$m \cdot t_{PCC} + \max_{i \in [1,\lfloor n \rfloor]}(s_i)$
t_R	$p_i \cdot t_{TTR} + (1-p_i) \cdot \infty$	$(1-(1-p_1)(1-p_2)...(1-p_m)) \cdot t_{TTR} + (1-p_i)(1-p_2)...(1-p_m) \cdot \infty$	$p_1 \cdot t_{TTR} + (1-p_1)p_2 \cdot (2 \cdot t_{TTR} + t_T) + ... + (1-p_1)...(1-p_{i-1})p_i \cdot (i \cdot t_{TTR} + (i-1) \cdot t_T) + ... + (1-p_i)(1-p_2)...(1-p_m) \cdot \infty$
d	t_T	$p_1 p_2 ... p_n \cdot m \cdot t_T$ + *dégradation non linéaire*	t_T

En comparant avec l'équation (3.3), on remarque qu'il n'y a aucune dégradation du délai à ce niveau. Le Tableau 3.1 résume les différents délais.

3.2.4 Étude du délai pour la classe C2

Pour la méthode de restauration IP, le principal élément du délai est le temps de convergence des tables de routage. En effet, tant que les tables de routage ne sont pas à jour, il est impossible de calculer les plus courts chemins et de rerouter le trafic. De plus, la méthode de restauration du niveau 3 avec OSPF et la méthode de restauration de niveau 3 proposée sont sensiblement identiques. De plus, il n'y a pas de différence ni sur les temps de transmission sur le chemin restauré, ni sur le temps d'établissement des LSPs. Ainsi, seule la différence entre les temps de restauration sera évaluée.

Dans la méthode de restauration de IP avec OSPF, le délai de restauration est composé de trois parties. La première est le temps de convergence des tables de routage. Cette étape prend généralement un temps de l'ordre de la minute. La deuxième étape consiste à calculer les plus courts chemins. L'utilisation d'OSPF permet de combiner le calcul des plus courts chemins avec la convergence des tables de routage. Le dernier élément est le temps de commutation du trafic. Ce temps est de l'ordre des millisecondes et sera négligé face au temps de convergence des tables de routage. On obtient alors l'équation suivante :

$$t_R = t_{OSPF} \tag{3.10}$$

On obtient alors un temps de restauration de l'ordre des minutes.

Dans la méthode de restauration de niveau 3 proposée, on ajoute une étape. Ainsi, on joint une composante au délai. Pour être certain d'obtenir les valeurs adéquates de bande passante disponible, la bande passante utilisée par le trafic de type C1 sera soustraite de la valeur de la bande passante des liens dans les métriques OSPF. Pour s'assurer d'avoir des valeurs adéquates pour les métriques, il faut donc attendre la fin de la restauration du trafic de type C1. Cette restauration, pour respecter les critères de qualité de service, prend un temps de l'ordre de 150 ms. Ce délai est négligeable face au

délai de convergence des tables de routage. Il est même probable que ce délai ne soit pas détecté puisque la fréquence d'échange des messages OSPF est de l'ordre des dizaines de secondes. En négligeant ce délai additionnel, on obtient :

$$t_R = t_{OSPF} \tag{3.11}$$

Ce qui est identique au délai encouru par la méthode de restauration de IP avec OSPF.

Dans le cas de la restauration de niveau 3, il n'y a pas de différence discernable entre les architectures proposées et la méthode traditionnelle de survivabilité au niveau du délai. On remarquera toutefois une différence au niveau de la quantité de trafic restaurée.

3.3 Étude de la survivabilité

Dans cette section, l'étude de la survivabilité des différentes méthodes est présentée pour les deux classes de trafic.

3.3.1 Méthode de commutation à l'ingress, classe C1

Avec la méthode de commutation à l'ingress, l'espérance mathématique de la quantité de trafic de voix restaurée est donnée par la somme de la bande passante de tous les LSPs de protection des chemins ayant subi une panne, pondérée par la probabilité de survie de chacun des LSPs de protection. Ceci nous donne l'équation suivante :

$$E(q) = \sum_{\forall LSPi} p_i b_i \tag{3.12}$$

Il est possible de développer la probabilité de survie de chacun des LSPs de protection. Cette probabilité correspond à la probabilité de la survie de tous les nœuds sur le chemin. En utilisant l'hypothèse d'indépendance de la survie des nœuds présentation dans la modélisation, il est possible de calculer la probabilité de survie du chemin en fonction de la probabilité de survie des noeuds Ainsi, si la probabilité de survie est la même pour tous les nœuds, on obtient :

$$E(q) = \sum_{\forall LSP} p_s^{|n|_{LSPi}} b_i \tag{3.13}$$

Cette équation sera utilisée pour comparer les performances des deux architectures proposées.

3.3.2 Méthode avec division de la bande, classe C1

Dans le cas de la méthode avec division de la bande passante, la quantité de trafic restaurée sera égale à la somme de la bande passante de chacun des chemins de secours ayant survécu étant donné que chacun des canaux est utilisé. L'espérance mathématique de cette quantité est la somme, sur tous les LSPs ayant subi une panne, de l'espérance de bande passante disponible sur les canaux de secours. Il est possible d'obtenir cette valeur en considérant la bande passante d'un seul canal pondéré par la probabilité de survie d'un canal unique additionné à la bande passante de deux canaux pondérée par la probabilité de survie de deux canaux sur *m*, et ainsi de suite. On obtient alors la formulation suivante :

$$E(q) = \sum_{\forall LSPi} \begin{pmatrix} \left(p_s^{|n|_{LSP1}} (1 - p_s^{|n|_{LSP2}}) \ldots (1 - p_s^{|n|_{LSPm}}) \frac{b_i}{m} \right) + \ldots \\ + \left(p_s^{|n|_{LSP1}} p_s^{|n|_{LSP2}} (1 - p_s^{|n|_{LSP3}}) \ldots (1 - p_s^{|n|_{LSPm}}) \frac{2b_i}{m} \right) \\ + \ldots \\ + \left(p_s^{|n|_{LSP1}} p_s^{|n|_{LSP2}} \ldots p_s^{|n|_{LSPm}} b_i \right) \end{pmatrix} \qquad (3.14)$$

Il est difficile de comparer cette valeur à l'espérance de quantité de bande passante rétablie par la protection par commutation à l'ingress. Toutefois, on remarque la présence d'états intermédiaires entre le meilleur cas (la protection de toute la bande passante) et le pire cas (la perte de la connexion). Ces états intermédiaires étant plus probable que les états extrêmes, la variance des résultats sera plus petite pour une topologie donnée. Ceci permet d'augmenter la prédictibilité de la survivabilité du réseau. Toutefois, il est difficile de visualiser le gain de bande passante sans utiliser de chiffres précis. La démonstration sera donc faite par simulation.

3.3.3 Méthode sans division de la bande, classe C1

Pour la méthode sans division de la bande, la quantité de trafic restaurée sera égale à la somme de la bande passante de chacun des chemins de travail ayant survécu à une panne. La probabilité de survie d'un chemin de travail est égale à la probabilité qu'au moins un de ses chemins de secours ait obtenu l'autorisation d'augmenter sa bande passante. On obtient alors la formulation suivante :

$$E(q) = \sum_{\forall LSPi} \begin{pmatrix} p_s^{|n|_{LSP1}}(1-p_s^{|n|_{LSP2}})...(1-p_s^{|n|_{LSPm}}) \\ +(1-p_s^{|n|_{LSP1}})p_s^{|n|_{LSP2}}(1-p_s^{|n|_{LSP3}})...(1-p_s^{|n|_{LSPm}}) \\ +...+p_s^{|n|_{LSP1}}p_s^{|n|_{LSP2}}(1-p_s^{|n|_{LSP3}})...(1-p_s^{|n|_{LSPm}}) \\ +...+p_s^{|n|_{LSP1}}p_s^{|n|_{LSP2}}...p_s^{|n|_{LSPm}} \end{pmatrix} b_i \qquad (3.15)$$

Cette formulation illustre bien que la survie de multiples chemins de secours n'augmente pas la quantité de trafic restaurée. Contrairement à la méthode avec division de bande, on remarque une augmentation de l'espérance de bande passante restaurée. Toutefois, pour faire une comparaison générale avec la méthode de protection par commutation, il faudrait considérer la topologie du réseau. En effet, un réseau dont la topologie augmenterait considérablement la longueur moyenne des chemins de protection (par rapport au LSP de secours de la méthode de protection par commutation) pourrait faire diminuer la probabilité de survie des chemins de secours et ainsi diminuer l'espérance mathématique de la bande passante restaurée. Si on prend l'exemple d'un LSP de bi = 2 Kbps protégé par 2 LSPs de secours ayant une probabilité de survie de 50%, l'espérance mathématique de la solution proposée est de 1,5 Kbps. Ceci représente une augmentation de 50% par rapport à une protection par commutation dont le LSP de secours possède une chance de survie de 50%.

Il est possible de résumer les différentes espérances mathématiques de la quantité de bande passante restaurée au Tableau 3.2.

Tableau 3.2 Comparaison de la survivabilité du trafic de type C1

Espérance de la quantité de bande passante restaurée	
Protection par commutation à l'ingress	$E(q) = \sum_{\forall LSP} p_s^{\lvert n \rvert_{LSPi}} b_i$
Architecture avec division de la bande passante	$E(q) = \sum_{\forall LSPi} \begin{pmatrix} \left(p_s^{\lvert n \rvert_{LSP1}} (1 - p_s^{\lvert n \rvert_{LSP2}}) ... (1 - p_s^{\lvert n \rvert_{LSPm}}) \frac{b_i}{m} \right) + ... \\ + \left(p_s^{\lvert n \rvert_{LSP1}} p_s^{\lvert n \rvert_{LSP2}} (1 - p_s^{\lvert n \rvert_{LSP3}}) ... (1 - p_s^{\lvert n \rvert_{LSPm}}) \frac{2b_i}{m} \right) \\ + ... \\ + \left(p_s^{\lvert n \rvert_{LSP1}} p_s^{\lvert n \rvert_{LSP2}} ... p_s^{\lvert n \rvert_{LSPm}} b_i \right) \end{pmatrix}$
Architecture sans division de la bande passante	$E(q) = \sum_{\forall LSPi} \begin{pmatrix} p_s^{\lvert n \rvert_{LSP1}} (1 - p_s^{\lvert n \rvert_{LSP2}}) ... (1 - p_s^{\lvert n \rvert_{LSPm}}) \\ + (1 - p_s^{\lvert n \rvert_{LSP1}}) p_s^{\lvert n \rvert_{LSP2}} (1 - p_s^{\lvert n \rvert_{LSP3}}) ... (1 - p_s^{\lvert n \rvert_{LSPm}}) \\ + ... + p_s^{\lvert n \rvert_{LSP1}} p_s^{\lvert n \rvert_{LSP2}} (1 - p_s^{\lvert n \rvert_{LSP3}}) ... (1 - p_s^{\lvert n \rvert_{LSPm}}) \\ + ... + p_s^{\lvert n \rvert_{LSP1}} p_s^{\lvert n \rvert_{LSP2}} ... p_s^{\lvert n \rvert_{LSPm}} \end{pmatrix} b_i$

3.3.4 Étude de la survivabilité pour la classe C2

Pour la classe de trafic C2, il est impossible de faire une évaluation analytique des gains de bande passante. L'avantage de la solution proposée est de permettre au trafic de classe C2 d'éviter les points de congestion causés par la restauration du trafic de type C1. Cette restauration risque de s'effectuer sur les liens avec une large bande passante puisque ceux-ci sont plus aptes à accorder des augmentations de bande passante. Or, les liens à large bande passante sont favorisés dans le calcul des métriques OSPF et se retrouvent donc sur le plus court chemin du trafic meilleur effort dans la méthode de

restauration de couche 3. Il y a donc une bonne possibilité que le trafic de type meilleur effort soit routé vers des points du réseau ne disposant pas assez de ressources pour assurer son expédition. Ceci peut créer la famine chez le trafic de type C2 puisque les ressources sont accordées en priorité au trafic des classes supérieures. La méthode suggérée permet d'éviter ces points de congestions. Ainsi, il est raisonnable de s'attendre à une augmentation de la quantité de trafic restauré. Toutefois, une démonstration analytique dépend fortement de la topologie du réseau, de la quantité de trafic dans le réseau, de la localisation des pannes et de la taille des liens.

3.4 Modélisation des architectures de protection

Cette section détaille le modèle des deux architectures de protection proposées. Tout d'abord, quelques généralités communes aux deux architectures de protection seront présentées. Ensuite, la modélisation de la méthode avec division de bande passante puis la modélisation de la méthode sans division de bande passante seront développées.

3.4.1 Caractéristiques communes des architectures de protection

Toutes les architectures de protection sont composées de trois éléments : un module de détection de panne, un module de notification de panne et un module de recouvrement lui-même divisé en deux sous-modules. Le premier sous-module est dédié au reroutage et le deuxième à l'allocation de ressources. L'architecture générale est présenté à la Figure 3.1.

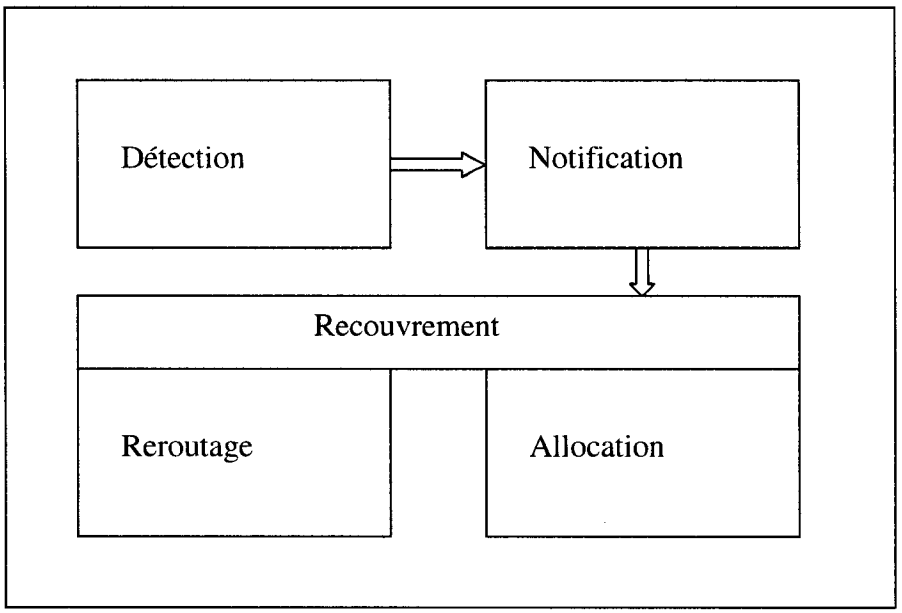

Figure 3.1 Architecture générale de survivabilité

Dans toutes les architectures de protection, on commence par détecter la panne et on envoie la notification de panne au point de réparation. Une fois le point de réparation notifié, les actions de recouvrement qui peuvent nécessiter l'utilisation des sous-modules sont lancées. L'ordre dans lequel sont utilisés les sous-modules dépend du type de mécanisme de survivabilité choisi. Dans un mécanisme utilisant la pré-allocation de ressources, le choix du chemin de secours (le reroutage) et la réservation des ressources du chemin de secours (allocation) sont effectuées avant la panne. Dans ce cas, le module de recouvrement ne fait que commuter le trafic. Dans un mécanisme utilisant la post-allocation, le reroutage et l'allocation sont effectués après la panne. Le module de recouvrement utilise premièrement le module de reroutage pour calculer les nouveaux chemins et ces chemins sont envoyés au module d'allocation pour faire la réservation sur ces chemins. Finalement, dans un mécanisme utilisant la pré-définition, on calcule les chemins avant la panne, mais la réservation est effectuée après la panne. Le module de recouvrement fait donc la commutation et utilise le sous-module d'allocation pour faire les réservations.

Pour faciliter le développement des architectures proposées, on énonce un certain nombre d'hypothèses de fonctionnement. Tout d'abord, on suppose un réseau de type métropolitain composé d'un cœur partiellement maillé et d'un réseau d'accès en étoile. Ceci représente une topologie communément utilisée dans les réseaux des grandes villes opérés par des intérêts privés. Deuxièmement, on suppose que le réseau devra supporter du trafic de multiples classes de services, incluant des classes avec des contraintes de qualité de service. Ce réseau utilisera IP sur MPLS avec le paradigme Diff-serv pour le traitement de la qualité de service. Ceci correspond à un réseau cellulaire de prochaine génération avec des usagers mobiles pouvant à la fois faire des appels téléphoniques, mais aussi avoir accès à Internet à partir de leurs équipements. La convergence vers un réseau tout-IP avec MPLS est aussi une caractéristique probable des réseaux de prochaine génération. Toujours dans un but de réalisme, on supposera l'utilisation d'OSPF comme protocole de transmission d'information de routage et l'utilisation de RSVP-TE comme protocole de signalisation pour MPLS. Il est ainsi possible de résumer les hypothèses comme suit :

1. Réseau métropolitain avec cœur partiellement maillé et accès en étoile;
2. Utilisation de IP/OSPF comme protocole de couche 3;
3. Utilisation de « Diff-serv aware » MPLS/RSVP-TE pour assurer la qualité de service.

Ces hypothèses permettent de détailler la plupart des modules de l'architecture de survivabilité.

Pour le module de détection des pannes, on utilise le même mécanisme dans tous les cas. En effet, l'utilisation des « hello messages » permet aux routeurs de détecter si certains de leurs voisins sont tombés en panne. Cette détection se fait à intervalle régulier avec augmentation de la fréquence dans les cas ou une panne est suspectée (par exemple, dans le cas où il y a une interruption de transmission). Pour la notification de panne, la couche IP et la couche MPLS n'utilisent pas le même mécanisme de notification. Dans le cas d'IP, la notification se fera par l'entremise d'OSPF qui, peu à peu, éliminera les routes brisées des tables de routage. Dans le cas de MPLS, le message

PathErr du protocole de signalisation RSVP-TE est utlisé. Les autres modules seront modifiés par les solutions proposées.

3.4.2 Architecture avec division de la bande passante

Dans l'architecture avec division de la bande passante, on divise le LSP de protection en plusieurs canaux. On utilise ensuite tous ces canaux pour transmettre le trafic de type C1. On utilise donc un mécanisme de pré-allocation pour le trafic possédant des requis de qualité de service. Une fois le trafic de type C1 restauré, on enclenche un mécanisme de protection de la couche IP. On utilise donc un mécanisme de post-allocation pour le trafic de type C2. L'architecture se divise donc en trois parties : une partie avant la panne et deux parties après la panne.

Puisqu'il s'agit d'une méthode de pré-allocation, il faut calculer les chemins de secours et allouer les ressources à ces chemins avant l'occurrence de la panne. Ceci représente la première partie de l'architecture. Pour chaque LSP qui nécessite une protection, on signale m chemins de secours (disjoints) à l'aide du protocole RSVP-TE. Le choix de ces chemins peut être basé sur une heuristique quelconque, mais le prototype utilisera les choix de l'opérateur comme heuristique pour déterminer ces chemins. La bande passante reservée sur chacun de ces chemins devra être de b/m Kbps, où b est la largeur de bande (en Kbps) du chemin de travail. Ces m chemins de secours deviendront les m canaux de secours après la panne.

Au moment de la panne, le module de restauration devra commuter le trafic sur le chemin de secours. Cette procédure s'effectue pareillement à une commutation de MPLS, à la seule différence que les paquets seront séquentiellement acheminés sur chacun des canaux, c'est-à-dire que le premier paquet sera acheminé sur le premier canal, le deuxième sur le deuxième canal et ainsi de suite. Les paquets pourront être réordonnancés à l'arrivée en utilisant une mémoire tampon. Au fur et à mesure que les pannes sur les canaux de secours sont détectées (après l'acheminement des messages *PathErr*), les canaux affectés sont retirés de la ronde de distribution.

Une fois la restauration des LSPs protégés terminée, la restauration de niveau 3 est enclenchée. La fin de la restauration de niveau 2,5 sera détectée par un chronomètre qui comptera une seconde après détection de la panne. Ce seuil est déduit de la norme de 150 ms comme délai maximal sur la voix. On amplifie ce seuil d'un ordre de grandeur (environ 10x) comme marge de sécurité. Une fois le chronomètre écoulé, les routeurs recalculeront leurs métriques OSPF ($10^8/d$ où d est le débit du lien en bps) en tenant compte de la bande passante réservée pour le trafic de voix. Ainsi, un lien dont la moitié de la bande passante est réservée pour du trafic de voix aura une métrique OSPF de $10^8/0{,}5d$. Avec cette modification, il est possible d'utiliser le mécanisme de reroutage IP pour restaurer le trafic de type C2.

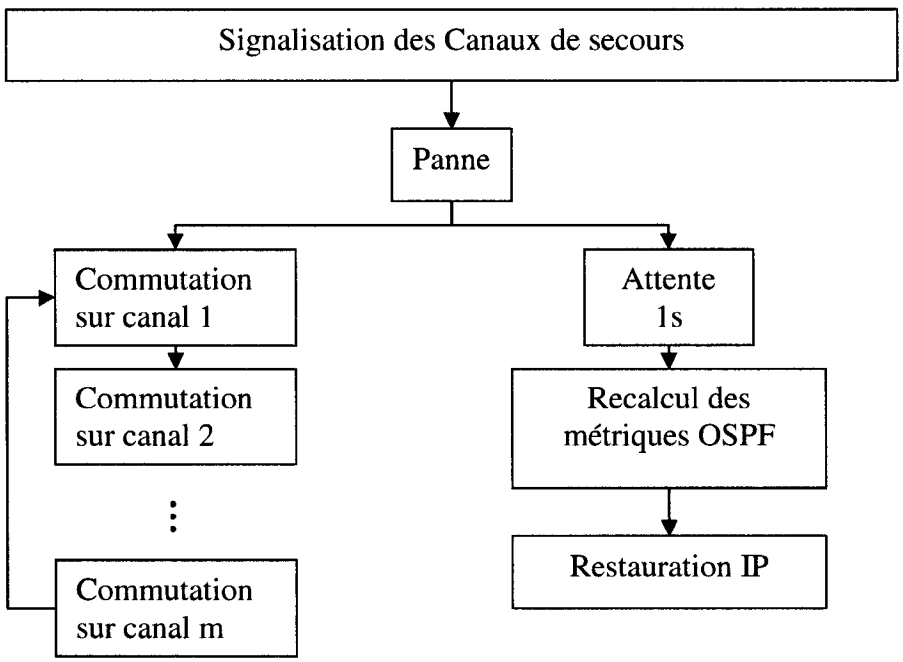

Figure 3.2 Mécanismes de survivabilité de la solution 1

La Figure 3.2 illustre les divers mécanismes utilisés dans l'algorithme avec division de bande passante.

3.4.3 Architecture sans division de la bande passante

Dans l'architecture sans division de la bande passante, on veut éviter la dégradation du délai causée par le fractionnement de la bande passante. La solution est donc sensiblement identique à l'architecture précédente, exception faite d'un mécanisme pour assurer l'utilisation d'un seul canal.

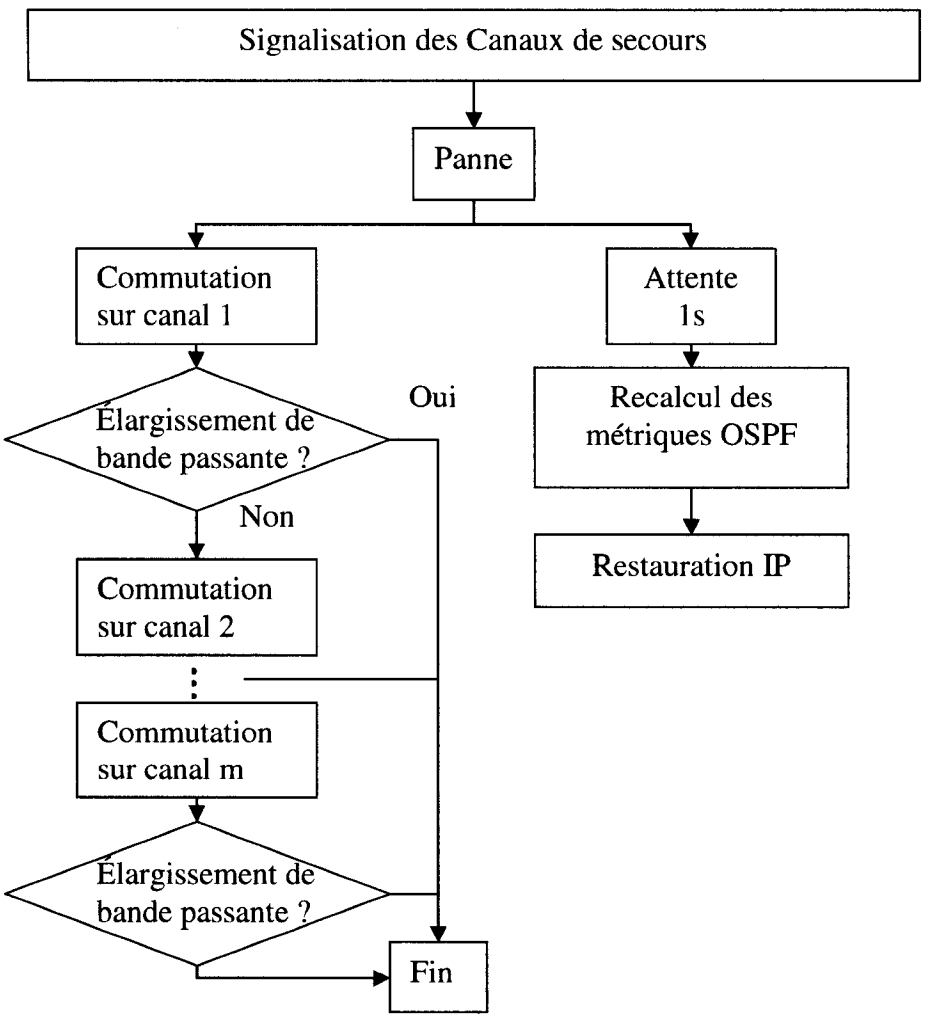

Figure 3.3 Mécanismes de survivabilité de la solution 2

Le seul module qu'il est nécessaire de modifier par rapport à la solution avec division de la bande passante est le module de commutation. Plutôt que d'envoyer une partie de la bande passante sur chaque canal, le module envoie la totalité de la bande passante sur un chemin choisi au hasard. Avec le message, le nœud d'origine (ingress) envoie une demande d'augmentation de la bande passante suivant la procédure décrite à la section 4.6.4 du RFC3209 (RSVP-TE : Extensions to RSVP for LSP tunnels). Dans le cas où la demande est acceptée, le nœud de destination (egress) répond en envoyant le message de réservation sur le canal utilisé et envoie un message *PathTear* pour désallouer la bande passante des autres canaux. Dans le cas où la demande est refusée, soit parce qu'il n'y a pas de bande passante disponible, soit parce qu'il y a une panne sur le chemin de secours, le canal est désalloué et on essaie avec le canal suivant. Lorsqu'il n'y a plus de canaux possibles, on coupe la connexion.

La Figure 3.3 illustre les divers mécanismes utilisés dans l'algorithme sans division de bande passante. Ces modèles des deux solutions proposées serviront à prototyper les solutions pour la simulation.

3.5 Prototypage

Le prototypage des architectures proposées vise à concrétiser le modèle analytique présenté aux sections précédentes. Cette section insiste sur l'environnement expérimental et les principaux éléments à considérer dans le prototypage.

L'environnement expérimental est un ordinateur de bureau Pentium IV de HP muni du logiciel de simulation OPNET Modeler 11.0.

Le prototype servira à déterminer la performance de chacune des solutions proposées. Pour établir la comparaison entre les deux solutions et les méthodes disponibles dans la littérature, on considérera principalement deux métriques : le délai et la quantité de trafic restaurée. Il faudra aussi conserver la priorité des transmissions. Finalement, il faudra considérer la coordination entre les mécanismes de protection des différents niveaux de la couche OSI.

La mesure du délai servira à évaluer la qualité du service fourni aux usagers. Le délai sera mesuré comme la différence de temps entre l'arrivée (au serveur) et le départ (du client) de chacun des paquets de l'application. Pour se trouver à l'intérieur de bornes acceptables, le délai devra être inférieur à 150 ms pour le trafic de type C1 et inférieur à 10 minutes pour le trafic de type C2 (c'est-à-dire que le délai doit être de l'ordre des minutes). Ce sont les délais maximaux suggérés pour préserver la qualité de service.

Bien que l'utilisation de la borne supérieure sur le délai à travers un seul AS puisse être considérée comme un mauvais indicateur de qualité de service, il est important de considérer que le réseau se trouve dans une situation de panne catastrophique. Dans ce type de situation, l'usager s'attend à une interruption de service. Ainsi, si le réseau peut fournir une connexion de qualité acceptable plutôt que supérieure, le réseau dépassera déjà les attentes de l'usager. Il serait donc possible de dégrader encore plus le délai, mais le prototype utilisera les valeurs suggérées comme seuil à partir duquel le service est trop dégradé pour être utile. Une étude plus approfondie des attentes des clients en présence de catastrophe n'entre pas dans le cadre de ce mémoire.

Une autre composante du délai qui devra être mesurée, est le délai de recouvrement. En effet, en mesurant le délai du départ de la source jusqu'à l'arrivée à la destination, on ne mesure pas le délai des paquets qui sont perdus. Puisque le trafic de type C1 laisse tomber les paquets dont le délai est trop grand, il serait possible de couper la connexion pendant un certain moment sans que l'effet ne se répercute sur le délai. Pour éviter les pertes excessives de paquets (qui se traduisent par une dégradation de la qualité de service), le temps de recouvrement sera mesuré. Le temps de recouvrement sera calculé comme la différence entre le temps d'émission du premier paquet après la panne qui possède une mesure de délai (inférieure à une borne acceptable) et le temps d'émission du dernier paquet avant la panne. Pour la classe C1, le temps de recouvrement devra être inférieur à 150 ms pour être considéré comme acceptable. Pour la classe C2, le temps de recouvrement devra être de l'ordre des minutes pour être considéré comme acceptable.

La deuxième mesure à considérer est la quantité de trafic restaurée. Cette quantité sera mesurée en Kbps. Cette métrique représente une quantification de la survivabilité du réseau. En effet, dans un réseau avec une mauvaise survivabilité, on remarquera la coupure de nombreuses connexions lors de l'occurrence d'une panne majeure. Le nombre élevé de coupure s'explique par le nombre élevé de pannes multiples coupant simultanément le chemin de travail et le chemin de secours. Pour s'assurer de la validité de la mesure, il faudra évaluer la quantité de trafic dont le délai peut être mesuré (c'est-à-dire qu'on ignore le trafic qui est transporté et perdu en cours de route) circulant sur le réseau en régime permanent. Il faut donc prendre la mesure après la fin de la restauration de niveau 2,5 pour déterminer la quantité de trafic de type C1 restauré, et après la fin de la restauration de niveau 3 pour déterminer la quantité totale.

Bien qu'il soit possible de restaurer une quantité maximale de trafic en ne restaurant que le trafic de type C2, il faudra porter une attention particulière aux priorités des classes de trafic. L'implémentation traitera quatre classes de trafic, soit :

1. Conversationnel mission critique (ex. appels 911);
2. Conversationnel (ex. appel téléphonique);
3. Meilleur effort assuré (ex. transactions financières);
4. Meilleur effort (ex. trafic Internet).

La classe 1 devra être prioritaire sur la classe 2, la classe 2 sur la classe 3 et ainsi de suite. Cette priorité devra permettre la préemption de classes inférieures dans le cas où le réseau ne dispose pas d'assez de ressources pour restaurer l'ensemble du trafic. Ceci permettra de simuler l'intelligence nécessaire au réseau pour prioriser les classes de trafic. Toutefois, dans un déploiement, l'utilisation de Diff-serv est préconisée (Diff-serv aware MPLS).

Un second problème de priorité réside dans la coordination entre les mécanismes de survivabilité des différents niveaux de la couche OSI. Nous avons vu au chapitre 2 qu'il pouvait y avoir compétition entre les différents niveaux, ce qui entraîne la multiplication des actions de protection à plusieurs niveaux simultanément. Toutefois,

dans les deux solutions proposées, nous n'avons pas à nous soucier des problèmes de coordination entre les niveaux. En effet, le niveau le plus bas utilisé dans les solutions est le niveau MPLS, ce qui correspond au niveau OSI 2,5. Il existe donc une possibilité de problème avec les niveaux 2 et 1. Or, aucun de ces deux niveaux ne possède l'intelligence nécessaire pour détecter les pannes de nœuds qui constituent les pannes les plus communes dans les cas de pannes majeures. Il est donc impossible d'entrer en conflit avec des mécanismes de survivabilité des niveaux inférieurs. De plus, les solutions proposées traitent de la coordination entre les niveaux 2,5 et 3, éliminant les problèmes entre ces niveaux. Finalement, puisque le niveau 3 est le niveau le plus élevé utilisé, la coordination entre le niveau 3 et les couches supérieures ne fera pas l'objet de ce travail, mais devrait être considéré par les développeurs des mécanismes de survivabilité de niveaux supérieurs.

En gardant en mémoire les principaux éléments à considérer dans le prototypage, soient les caractéristiques et seuils du délai et de la survivabilité des solutions, ainsi que la priorité de certaines classes de trafic, nous pourrons déterminer les métriques pertinentes à l'analyse des solutions. De plus, les détails du modèle et de l'architecture permettront d'implémenter les architectures proposées sous l'environnement de simulation OPNET. Le détail de ces sections et l'analyse des résultats trouvés à partir de ces prototypes seront présentés au chapitre 4.

CHAPITRE IV

ANALYSE DES RÉSULTATS

Dans ce chapitre, nous présenterons les principaux résultats obtenus au cours des expériences. Dans un premier temps, nous présenterons la méthodologie utilisée pour obtenir les résultats, ensuite, nous présenterons les résultats obtenus à partir du modèle de simulation et nous terminerons en présentant les résultats obtenus à partir du modèle analytique.

4.1 Plan d'expérience

Afin de prouver la validité du modèle présenté, des simulations devraient être réalisées. Cette section décrit la démarche expérimentale qui devrait être suivie lors de ces expériences. Nous présenterons tout d'abord les objectifs et motivations du choix d'expérience; ensuite, nous présenterons les principales métriques utilisées. Enfin, nous détaillerons les tests qui seront effectués.

4.1.1 Objectifs

Lors de cette phase de simulation, l'objectif principal est de faire la preuve de concept des solutions proposées afin de montrer la validité des solutions avancées. Puisqu'il ne s'agit pas d'une phase de développement, il n'est pas nécessaire d'implanter l'ensemble des fonctionnalités dans un prototype. Ceci motive l'utilisation de simulations pour faire la preuve de concept.

L'objectif des solutions présentées au chapitre 3 est d'améliorer la survivabilité des réseaux de télécommunications aux pannes catastrophiques. Pour démontrer l'amélioration, nous devrons comparer les résultats avec des mécanismes utilisés dans les réseaux de télécommunications contemporains. Afin de pouvoir tirer des conclusions significatives, nous utiliserons la méthode de protection par commutation à l'ingress par

MPLS comme référence. En effet, puisque les solutions proposées se basent sur une protection globale, il est nécessaire de comparer avec une méthode de protection globale pour éviter que ce choix d'implémentation n'affecte les résultats. Il aurait été possible de faire une implantation locale des solutions proposées et de comparer les résultats avec MPLS FAST-REROUTE, mais l'implémentation d'une méthode locale est plus complexe et nullement nécessaire pour faire la preuve de concept. De plus, cette référence est considérée comme un des mécanismes les plus performants sur le marché au moment de l'écriture de ce mémoire. Ainsi, en se comparant avec cette référence, il n'est pas nécessaire de se comparer avec d'autres mécanismes moins performants.

Aussi, étant donné la similarité des méthodes et les limitations du simulateur utilisé, seule la proposition 1 sera testée. Ce choix peut être justifié aisément en comparant les performances attendues présentées au chapitre 3. En effet, les analyses mathématiques indiquent que la proposition 2 sera plus performante que la proposition 1. Ainsi, une preuve de concept de la proposition 1 implique la validité de la solution 2. Ceci nous permet d'éviter plusieurs complications au niveau du simulateur.

4.1.2 Métriques

Les principales métriques de performance en réseautique sont la disponibilité des ressources, le délai, la gigue et la perte de paquet. Ces métriques, principalement reliées à la qualité de service, ne sont pas toutes adaptées à une étude de survivabilité.

La métrique de disponibilité des ressources est la facette la plus importante dans le cadre d'une étude de survivabilité. En effet, l'objectif de la survivabilité est d'assurer un service même en cas de panne. La disponibilité de ces services est directement liée à la disponibilité des ressources. Toutefois, l'estimateur traditionnel de la disponibilité des ressources (le taux de blocage) est mal adapté à notre étude. En effet, dans un cas de panne catastrophique, toutes les ressources risquent d'être utilisées, ce qui rendra le taux de blocage très grand et moins significatif. Pour obtenir une meilleure évaluation dans le cadre d'une simulation par OPNET, nous estimerons cette métrique par la quantité de

trafic restaurée, en faisant l'hypothèse que toutes les ressources auront été utilisées pour restaurer du trafic. Nous détaillerons la méthode pour mesurer cette métrique plus loin.

La métrique de délai est aussi une métrique importante dans le cadre d'un scénario de panne catastrophique. En effet, lorsqu'on se retrouve avec des classes de trafic sensibles aux délais, il est possible de dégrader la qualité de service au delà d'un seuil acceptable pour la communication. Après ce seuil, il sera inutile de continuer la communication et ce trafic ne devrait pas être considéré comme du trafic restauré. Le délai devient donc un indicateur important pour déterminer la quantité de trafic restaurée. La façon d'estimer cette métrique sera détaillée plus tard.

Pour les deux autres métriques, soit la gigue et la perte de paquets, ces métriques sont moins intéressantes dans le cadre d'une étude portant sur les pannes catastrophiques. En effet, lorsqu'une panne survient (et les effets sont encore plus marqués s'il s'agit d'une panne majeure), on observe assurément la perte des paquets en transit, ce qui cause une augmentation incroyable de la gigue dans les moments qui suivent une panne. L'envergure de cette dégradation pour un scénario de panne catastrophique est énorme, ce qui nous pousse à ne pas considérer ces statistiques précises dans le cadre de notre étude.

Pour estimer le délai, la statistique de délai de bout en bout (tel que vu par l'application) sera utilisée. Cette statistique est le meilleur indicateur de la qualité de service perçue par l'usager. Bien qu'il y ait un certain nombre de métriques additionnelles de qualité de service, elles ne sont pas jugées pertinentes dans le cadre de cette recherche. Mentionnons principalement que la gigue et le blocage ne pourront être évités dans un cas de panne catastrophique et que les résultats provenant de ces métriques seraient donc peu significatives.

Pour estimer la quantité de trafic restaurée, nous utiliserons une somme de statistiques. La quantité de trafic restaurée sera la somme de la largeur de bande consommée par l'ensemble des trafics dont le délai est inférieur à 300 ms. Au-delà de ce seuil, les communications sont jugées par l'expérimentateur comme rompues. Ce seuil est une valeur arbitraire basée sur le double du requis minimal en mode de

fonctionnement normal. Le raffinement de ce seuil sera laissé à d'autres chercheurs et aucune justification ne sera présentée dans ce mémoire autre que la nécessité. Ce seuil exclut implicitement toutes les classes sans requis de qualité de service puisque la survivabilité de ces classes aux pannes catastrophique est assurée en autant que la connectivité IP est préservée. Ainsi, la quantité de trafic sans requis de qualité de service restaurée sera la même avec toutes les méthodes. Cette information n'est pas jugée pertinente dans le cadre d'une comparaison.

4.1.3 Tests optimaux

Pour couvrir le plus grand nombre de cas possibles, nous devrions effectuer quatre séries de tests. Les trois premiers tests seront basés sur les principales variables contrôlées par l'utilisateur, soient la probabilité de survie des nœuds, la topologie du réseau et la quantité de trafic circulant dans le réseau. Le quatrième test sera plutôt basé sur la distribution spatiale des pannes.

Dans la première série de tests, nous comparerions les performances de la protection par commutation à l'ingress de MPLS et de la solution proposée 1 en fonction de la probabilité de survie des nœuds. Pendant l'exécution de ces tests, nous utiliserions une topologie et une quantité de trafic constantes. La probabilité de survie des nœuds sera variée entre 25 et 75 %. Le cas où la probabilité de survie est de 75% est considéré comme le pire cas pour notre solution puisque la probabilité de panne multiple (sur le chemin de travail et le chemin de secours) est plus faible. Ainsi, les mécanismes traditionnels de protection devraient répondre adéquatement à la panne.

Dans la deuxième série de tests, nous garderions la probabilité de panne et la quantité de trafic dans le réseau constantes, mais nous ferions varier la topologie. Nous choisirons plusieurs topologies jugées typiques, mais toutes ces topologies devront avoir un minimum de diversité pour permettre au minimum deux chemins de secours disjoints.

Dans la troisième série de tests, nous conserverions la même topologie et une probabilité de survie des nœuds constante, mais nous ferons varier la quantité de trafic dans le réseau. Nous ferions varier le taux d'utilisation de 50 à 98 % de la capacité totale

du réseau. Le cas d'utilisation à 50% sera considéré comme le pire cas pour notre solution puisque la probabilité qu'une grande quantité de bande passante soit disponible après la panne est non négligeable.

Finalement, pour la quatrième série de tests, nous fixerions la topologie et la quantité de trafic, mais nous ferions varier la distribution spatiale des nœuds qui tombent en panne (éliminant ainsi le paramètre de probabilité de survie des nœuds). Ces tests pourraient déterminer la validité de la solution par rapport à divers scénarios de pannes corrélées qui ne sont pas nécessairement reliées à des catastrophes naturelles. Mentionnons à titre d'exemple les pannes causées par une attaque de déni de service. Ces attaques provoquent généralement des pannes localisées autour de l'élément attaqué ou sur le chemin privilégié par le trafic hostile. Ces tests permettraient d'élargir le champ de validité de la solution proposée.

Malheureusement, compte tenu des limitations inhérentes aux simulateurs disponibles, il est impossible d'implémenter ce plan d'expérience. En effet, OPNET (qui représente le simulateur le plus complet disponible sur le marché au moment de l'écriture de ce mémoire) rend nécessaire le chargement des liens à une valeur très près de la capacité. À cause de cette nécessité de dimensionnement très précis, il est impossible d'ajouter beaucoup de variabilité aux expériences. Nous proposons donc une expérience hybride de simulation et de résultats analytiques en espérant que les développements des outils de simulation permettent un jour la mise en œuvre du plan d'expérience tel que proposé.

4.1.4 Tests réalisés

Dans un premier temps, un modèle de simulation est développé à partir du logiciel OPNET. Ce modèle, compte tenu des limitations d'OPNET, ne concerne que la solution avec division de la bande passante pour du trafic de voix. En utilisant une configuration particulière du réseau pour obtenir un chargement intéressant et une configuration de panne jugée catastrophique (non triviale), nous obtiendrons une comparaison de la performance de notre solution par rapport à la performance de la

commutation de protection à l'ingress par MPLS. Cette comparaison nous permettra d'affirmer qu'il est possible pour notre solution de performer aussi bien que MPLS, tel que prévu dans les modèles analytiques proposés au chapitre 3. Nous utiliserons ensuite ces modèles analytiques pour généraliser l'application de nos solutions.

La généralisation se fera en deux étapes. Les deux étapes utiliseront l'environnement MATLAB release 12 pour générer les résultats. Dans un premier temps, nous testerons l'effet sur la quantité de trafic restauré de faire varier les paramètres de probabilité de panne des nœuds, du nombre de nœuds par chemin de protection et du nombre de chemins de protection pour chacune de nos solutions. Ces courbes seront obtenues directement à partir des équations du chapitre 3. Dans un deuxième temps, toujours en utilisant le modèle mathématique du chapitre 3, nous générerons des résultats pour obtenir une étude statistique des résultats. Dans cette étude statistique, nous remplaçons l'espérance mathématique des équations du chapitre 3 par un générateur de nombres aléatoires que nous comparons avec la probabilité de panne. On peut ainsi obtenir des valeurs sur la moyenne et la variance de la quantité de trafic restaurée.

4.2 Résultats de simulation

Cette section détaille le modèle utilisé pour la simulation. En premier lieu, des limitations relatives à l'environnement de simulation et les hypothèses simplificatrices seront présentées. Ensuite, le modèle de simulation sera détaillé et finalement les résultats de simulation seront montrés.

4.2.1 Environnement de simulation OPNET

OPNET est le simulateur le plus reconnu dans le milieu académique au moment de l'écriture de ce mémoire. Cette réputation est tout à fait justifiée puisque OPNET permet de simuler rapidement, à l'aide d'une interface graphique, des réseaux complexes incluant une variété de protocoles et de technologies. Avec la version 11.0, beaucoup de

développements, notamment au niveau des réseaux ad hoc et de la mobilité ont été introduits. Toutefois, OPNET possède quand même plusieurs limitations qui nous contraignent à adapter le modèle.

Une première limitation importante de OPNET est l'absence de réservation dynamique. En effet, le protocole RSVP-TE n'est implémenté que dans le but de faire la signalisation des LSPs. Lorsqu'un message de réservation traverse le réseau, les différents nœuds du réseau retire la valeur de bande passante demandée d'une variable gardant en mémoire la quantité de ressources disponibles, mais n'assigne pas d'espace tampon ni ne change son ordonnancement de file d'attente d'aucune façon. Similairement, lorsqu'un nœud reçoit un message *PathTear*, il ajoute la valeur de bande passante du chemin à la valeur disponible dans le nœud sans affecter les files d'attentes. La fonction de réservation de RSVP-TE ne sert donc qu'à déterminer si la connexion est acceptée ou rejetée. Il devient donc très difficile d'implanter les solutions proposées qui sont basées en grande partie sur la réservation de bande passante. Heureusement, nous disposons de l'architecture *Diff-Serv* qui nous permet de faire de la priorisation.

Cette priorisation devra respecter la discipline de files d'attente stipulant que le trafic dont la bande passante est réservée devra toujours passer devant le trafic sans réservation. Puisque la variation dynamique des poids *Diff-Serv* n'est pas nous plus implantée, pour simuler cette discipline, nous utilisons une architecture à deux files d'attente : la file à faible délai (Low Lentency Queue ou LLQ) et la file par défaut (Default Queue ou DQ). Le trafic protégé sera affecté à des LSPs assignés à la LLQ tandis que le reste du trafic sera assigné à la DQ. Une politique de surveillance sur les LSPs sera utilisée pour remarquer le trafic dépassant la réservation à une classe de service inférieure passant par la DQ. Cette façon de fonctionner nous force à faire une première simplification importante : les simulations ne traiteront que d'un seul type de trafic. Compte tenu de la priorité relative des différentes classes de trafic, on doit utiliser deux files d'attente par type de trafic (une file pour le trafic protégé et une autre pour le trafic non-protégé). Puisque nous ne disposons que de deux files d'attente (LLQ et DQ), nous ne pouvons donc simuler qu'un seul type de trafic. Il serait possible d'inclure tous

les types de trafic en utilisant plusieurs queues avec différents poids, mais ce type d'expérience testerait les performances des poids choisis beaucoup plus que les performances des solutions proposées. Aussi, puisque la différence entre la solution 1 et la solution 2 se base principalement sur la variation de la méthode de réservation (avec ou sans division de la bande passante), seule la solution 1 (plus facile à implémenter) sera testée. Puisque cette solution est la solution qui présente les moins bonnes performances, elle sera jugée comme une borne inférieure de la performance de la deuxième solution.

Une deuxième limitation, moins sévère, est l'impossibilité de combiner des pannes de nœuds avec le module OSPF. Ceci est dû à une erreur dans le logiciel qui est en cours de révision par les développeurs du simulateur. Il a donc été nécessaire de remplacer les pannes de nœuds par la panne de tous les liens attachés à ces nœuds. Ceci est mathématiquement équivalent à la panne de nœud et cause peu de problème à l'analyse de performance. Toutefois, ceci augmente considérablement le temps nécessaire à l'adaptation des modèles pour faire différents scénarios et pourrait causer problème si les pannes n'étaient pas de durée infinie. En effet, puisque le nœud est encore en fonction, il peut garder l'information dans ses mémoires tampons (qui ne se remplissent jamais puisque l'accès au nœud est bloqué). Cette information, si l'accès est rétabli, pourrait retourner dans le réseau (même si elle n'est plus à date, les temps d'attente pour la retransmission étant dépassés) et fausser les résultats. Nous ne testerons donc pas l'impact de la restauration des nœuds.

Une troisième limitation, la plus importante, est une limitation provenant de la modélisation des files d'attente incluant la qualité de service. Cette file d'attente possède un certain nombre de paramètres dont les plus important sont la longueur maximale du tampon (en paquet), la longueur maximale des files de chacune des classes de qualité de service (ici LLQ et DQ, aussi en paquet) et la dimension de l'interface (en Kbps). La longueur maximale du tampon représente le nombre maximal de paquets pouvant être en attente sur cette interface. La longueur maximale des files d'attente représente le nombre maximal de paquets de chaque classe de service pouvant être en

attente. La dimension de l'interface est un indicateur du taux de service de la file d'attente. Tant que la longueur maximale du tampon n'est pas excédée, les différentes files d'attente peuvent mettre en attente un nombre arbitraire de paquets (c'est-à-dire que les contraintes sur le nombre de paquets maximal par file d'attente sont relaxées). Ceci correspond à dire que les mécanismes d'ordonnancement des files d'attente se font vraiment remarquer seulement lorsqu'on commence à dépasser la taille maximale du tampon. Parallèlement, si le simulateur juge que l'interface est congestionnée (c'est-à-dire que les paquets dans le tampon sont supérieurs à la capacité de traitement, donnée par la valeur de la dimension de l'interface), l'interface est tout simplement bloquée et les mécanismes de qualité de service ne s'appliquent plus. Il faut donc dimensionner le réseau de façon à ce que l'interface soit suffisamment congestionnée pour avoir la classification de paquets, mais pas trop pour éviter la congestion.

De plus, à partir de cette modélisation, il est théoriquement possible d'observer une oscillation de l'interface. En effet, s'il y a un débalancement sur la quantité de trafic allant dans les deux files d'attente (LLQ et DQ), à la limite de la taille du tampon, une interface aura largement dépassée sa limite, alors que l'autre sera bien en dessous de sa limite. Une fois la limite dépassée, la file surchargée se débarrassera d'un très grand nombre de paquets, ce qui fera tomber le nombre de paquets total dans le tampon bien en dessous de la taille maximale du tampon. Les différentes files d'attente retrouveront donc la possibilité de mettre en attente un nombre arbitraire de paquets.

Cette limitation rend donc un peu chaotique la gestion des files dans le réseau. Puisque la simulation d'une panne catastrophique suppose que la quantité de ressources est insuffisante pour restaurer tout le trafic, il est évident qu'on devrait se trouver en situation de congestion. Or, ce cas est impossible à simuler avec OPNET. Il nous est tout aussi impossible d'extraire des données pertinentes d'une simulation dans un réseau peu chargé puisque la surabondance de ressources atténue trop l'effet bonifiant d'avoir un LSP de protection et il est difficile d'observer une distinction claire entre le trafic protégé et le trafic non protégé. Il faut donc simuler un cas pour lequel on se trouve à l'extrême limite de la congestion. Ceci nous empêche de faire plus d'une expérience en

faisant varier les paramètres, puisque le dimensionnement du réseau doit être adapté de façon très serrée au scénario en cours. Aussi, comme les files d'attente sont ultrasensibles aux variations, une simple variation de l'assignation des chemins de protection peut faire varier l'utilisation des liens critiques du réseau de 20% (avant et après la panne). Ceci rend absolument impossible la comparaison de notre solution et de MPLS en terme de délais. Cette métrique a donc tout simplement due être abandonnée.

Compte tenu de ces limitations, le modèle de simulation servira à démontrer qu'il est possible d'obtenir les résultats (en terme de trafic restauré) prédits dans le modèle développé au chapitre 3. Cette démonstration ne se fera que pour la solution 1. La généralisation de l'analyse de performance s'effectuera à partir du modèle analytique à la section 4.3.

4.2.2 Modèle de simulation

Le modèle de simulation sera basé sur un réseau métropolitain hypothétique obtenu à partir d'une carte des lignes de transmission de données de la ville de New York. Il est réalisé sur OPNET version 11.0 sur un ordinateur de bureau Pentium 4 de HP.

La topologie du réseau est partiellement maillée. Le degré minimal des nœuds du réseau est de 3 et le degré maximal est de 4. Les sources de trafic et les nœuds d'agrégation se trouvent en périphérie et sont disposés en étoile. Une représentation graphique de la topologie utilisée est incluse à l'annexe 3. Les demandes de trafic ont été disposées afin que le trafic traverse la plus grande partie possible du réseau. Par exemple, les sources de trafic identifiées comme les nœuds 17 à 22 doivent se rendre au serveur identifié comme nœud 51 et les nœuds 41 à 46 doivent se rendre au nœud 50. Le déplacement d'un nœud d'agrégation dans le sens des aiguilles d'une montre pour les sources entraîne le déplacement d'un nœud d'agrégation dans le sens des aiguilles d'une montre pour le serveur. Les LSPs sont répartis de façon géographique de sorte que les

chemins de travail soient disjoints de nœud. Ceci évite la production d'un scénario aberrant où tout tomberait en panne ou encore tout serait restauré.

Pour le trafic, nous avons utilisé les générateurs de trafic de voix de qualité GSM fournis par OPNET. Seuls les attributs relatifs à la qualité de service (spécifiquement le code Diff-Serv lui étant associé) et les noms de serveurs utilisés ont été modifiés. Ce type de trafic a été choisi en raison du fait qu'il est à débit constant. En effet, GSM envoie des paquets, toujours de même taille, à intervalle régulier. Ceci facilite grandement le dimensionnement des liens puisque le débit est égal au débit binaire. Il est donc inutile de faire des calculs pour tenir compte de la taille maximale des rafales, du taux d'arrivée des paquets, etc. Un autre avantage de ce type de trafic est que le temps inter-arrivées des paquets est constant et inférieur à 100 ms. Ceci permet d'observer avec plus de précision l'effet des pannes qui peuvent se produire et se réparer dans un temps de silence avec un modèle probabiliste.

Le scénario de panne a été choisi pour faciliter l'observation de la restauration. Nous faisons donc tomber en panne les liens attachés aux quatre LSRs de la partie sud-est du réseau. Donc, en observant les sources 17 à 22, on se retrouve avec un seul lien entre l'origine et la destination. Toutes les autres paires origine-destination ne possèdent plus de chemins fonctionnels. Cette configuration de panne nous permet d'observer les trois cas possibles pour MPLS : un groupe avec un chemin de travail qui ne tombe pas en panne, un groupe avec un chemin de travail qui tombe en panne et dont le chemin de protection est disponible, un groupe dont le chemin de travail et le chemin de protection tombent en panne simultanément.

Chaque groupe de 6 nœuds est assigné à 3 chemins de travail. Pour le groupe 17 à 22 les nœuds 17 et 18 sont assignés au LSP_1, les nœuds 19 et 20 au LSP_2, et les nœuds 21 et 22 au LSP_3. Pour le modèle basé sur MPLS, chacun de ces LSPs est assigné à un LSP de protection (LSP_1 est protégé par le LSP_4, le LSP_2 par le LSP_5 et le LSP_3 par le LPS_6). Pour notre solution, on désagrège les groupes et chacun des nœuds possède un LSP de protection différent (nœud 17 est protégé par LSP_4, nœud 18 par LSP_5 et ainsi de suite) pour simuler la division de la bande passante.

À partir de cette topologie, on compare les statistiques du nombre de paquets sortant des LSPs pour comparer la quantité de trafic restaurée. Cette mesure est équivalente à une mesure de la quantité de trafic restaurée puisque les paquets du générateur GSM ont un débit et une taille constantes et puisque seul le trafic passant dans les LSPs entre dans la LLQ, lui assurant un délai convenable.

4.3.3 Résultats

Nous présenterons les résultats obtenus aux Figures 4.1 et 4.2. Ces résultats détaillent le trafic dans les LSPs pour les LSPs 1 à 9.

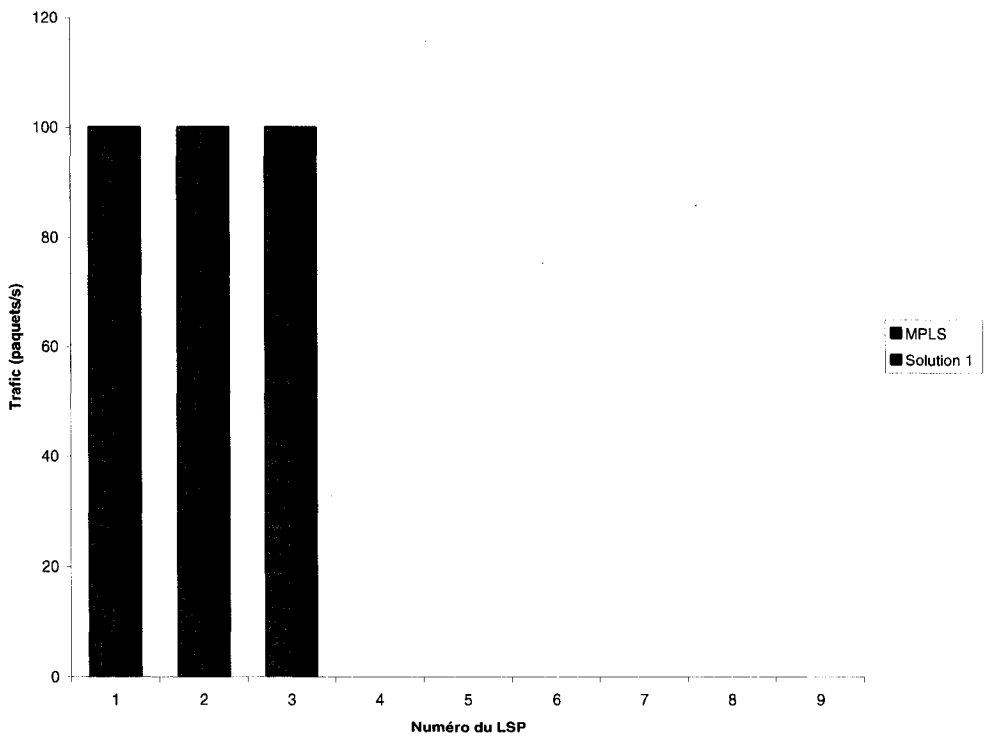

Figure 4.1 Trafic dans les LSPs avant la panne

On remarque qu'avant la panne, la situation est identique. Tout le trafic traverse les chemins de travail (LSP_1 à LSP_3). De plus, la quantité totale de trafic est de 300

paquets par seconde, ce qui servira de référence pour la quantité de trafic après la panne, présentée à la Figure 4.2.

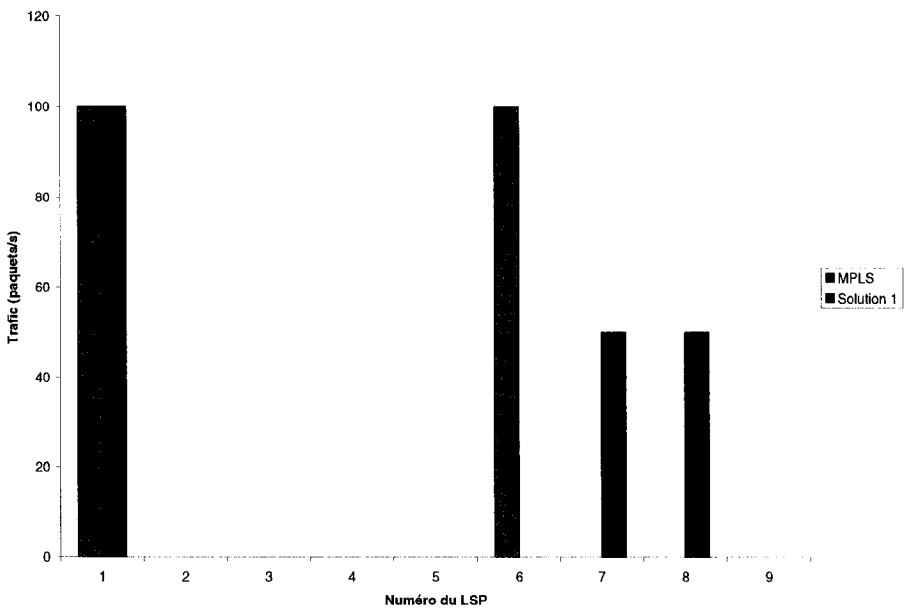

Figure 4.2 Trafic dans les LSPs après la panne

Cette figure montre que le chemin du LSP_1 ne tombe pas en panne (expliquant le débit identique de 100 paquets par seconde pour MPLS et la solution 1, ainsi que l'absence de trafic dans le LSP_4 qui est le LSP de protection pour MPLS et les LSPs 4 et 5 qui sont les LSPs de protection pour la solution 1). On voit aussi que le chemin du LSP_2 tombe en panne et que le LSP de protection MPLS (LSP_5) tombe aussi en panne. Par contre, malgré que le LSP de protection LSP_5 tombe en panne pour la solution 1, on sauvegarde tout de même 50% du débit avec le LSP_7. Finalement, on voit que le LSP de travail LSP_3 tombe en panne, mais est recouvert à 100% par MPLS (au LSP_6) et à 50% par la solution 1 (avec le LSP_8).

En analysant ces données, il est possible de tirer certaines conclusions. Tout d'abord, on remarque que, dans chacun des cas, la proportion de LSPs en panne est la même (4/6 pour MPLS et 6/9 pour la solution 1). Aussi, on peut remarquer que la quantité de trafic circulant après la panne est la même dans les deux cas (200 paquets par

seconde) et, par le fait même, que la proportion de trafic restaurée est identique (100 paquets par seconde sur 200, soit 50%). On peut donc déjà affirmer qu'il est possible pour notre solution la moins performante d'égaler la performance de la protection par commutation de MPLS. Ceci valide le fonctionnement des solutions proposées. Toutefois, il est possible d'extraire quelques informations supplémentaires de ces données.

Une première observation importante est que la méthode de protection par MPLS est une solution tout ou rien. Ainsi, on se retrouve avec un LSP rétabli à 100% tandis que l'autre est rétabli à 0%. C'est ce résultat qui nous donne une moyenne nette de 50% de restauration. Nous obtenons alors deux LSPs qui ne subissent pas de panne visible et un LSP qui subit une panne entraînant la coupure de la connexion. Avec notre approche, les deux LSPs de travail subissant une panne sont rétablis à 50%, pour une moyenne nette de 50%. On obtient alors un LSP qui ne subit pas de panne visible, deux LSPs subissant une dégradation de la qualité de service et aucun LSP ne subissant une perte de connexion. On obtient donc une restauration plus équitable.

Ainsi, la comparaison avec MPLS pour cette configuration particulière de réseau est encourageante. Afin de fournir une plus grande applicabilité de ces résultats, nous allons généraliser à partir du modèle analytique.

4.3 Modèle analytique

Le modèle analytique a été dérivé des équations résumées au Tableau 3.2 du chapitre 3. La première partie des tests effectués à partir de ce modèle mathématique utilise directement ces équations pour générer les résultats numériques. Dans la deuxième partie, la même modélisation est utilisée mais on substitue l'espérance mathématique par une série de générateurs aléatoires qui sont comparés aux seuils de pannes. On présente ensuite les résultats pour la moyenne et la variance de la quantité de trafic restaurée pour les classes de service à grande demande de qualité de service.

Pour générer les nombres aléatoires, la fonction *rand( )* de MATLAB release 12 a été utilisée. Cette fonction appelle un générateur de nombres pseudo-aléatoires fonctionnant à partir d'une machine à états. À chaque fois qu'un nouveau nombre aléatoire est demandé, la machine saute à l'état suivant pour éviter les répétitions. Cette machine contient 2^{1492} et peut donc générer 2^{1492} nombres avant de subir des répétitions. Toutefois, à chaque fois que l'application de MATLAB est redémarrée, le générateur est redémarré [26]. Aucune précaution n'a été prise à ce sujet, mais ne connaissant pas à priori la suite de valeurs générés par *rand( ),* nous estimons que ce désavantage a peu d'impact sur les résultats.

Outre cette limitation mineure de MATLAB, il a été nécessaire de poser quelques hypothèses simplificatrices pour faciliter l'implémentation des modèles mathématiques. Premièrement, afin de pouvoir faire des factorisations et produire un grand nombre de résultats de façon automatisée, il a été nécessaire de poser l'hypothèse des chemins de même longueur. La valeur de $|n|$ est donc considérée la même pour tous les LSPs d'une expérience. Deuxièmement, afin de pouvoir comparer facilement la solution 2 avec la méthode de protection par commutation de MPLS, il était nécessaire de garder la même paramétrisation pour la probabilité de panne. Ainsi, l'hypothèse stipulant que la probabilité de rejet d'une demande d'élargissement de bande passante est négligeable a été posée. Cette hypothèse est jugée comme ayant peu d'effet sur les résultats puisque les demandes d'élargissement de bande passantes seront rejetées, surtout lorsque la restauration aura atteint la capacité maximale du réseau (c'est-à-dire qu'il ne reste plus de ressources réseau disponible).

4.3.1 Modèle basé sur l'espérance mathématique

Dans cette expérience, on calcule l'expérience mathématique de la quantité de trafic restaurée. Cette quantité est la proportion de trafic restauré sur 10 LSPs (ainsi, une valeur de 1 correspond au rétablissement de la capacité totale des 10 LSPs et une valeur de 0.5 correspond à 50% de la capacité de ces 10 LSPs). Cette quantité est représentée en fonction de la probabilité de panne de nœuds (c'est-à-dire l'inverse de la probabilité

de survie). On montre les graphiques pour différentes longueurs de chemin de secours (c'est-à-dire pour différentes longueurs de $|n|$) ainsi que pour différentes valeurs du nombre de LSPs de secours (c'est-à-dire différentes valeurs de m). On présentera successivement les graphiques relatifs à la solution 1 pour des valeurs de $|n|$ de 2, 3, 5 et 10 et de la solution 2 pour les mêmes valeurs de $|n|$. Sur chacun de ces graphiques, on retrouvera l'espérance de la quantité de trafic restaurée pour la commutation de protection par MPLS (en vert), la solution courante avec deux chemins de protection (m=2) en rouge et la solution courante avec trois chemins de protection (m=3) en bleu.

Pour la solution 1, on remarque aisément que l'espérance mathématique de la quantité de trafic est égale à celle de MPLS. Aussi, on remarque que le nombre de chemins de secours n'a pas d'effet sur l'espérance de la quantité de trafic restaurée comme en témoignent les trois courbes confondues des figures 4.3 à 4.6.

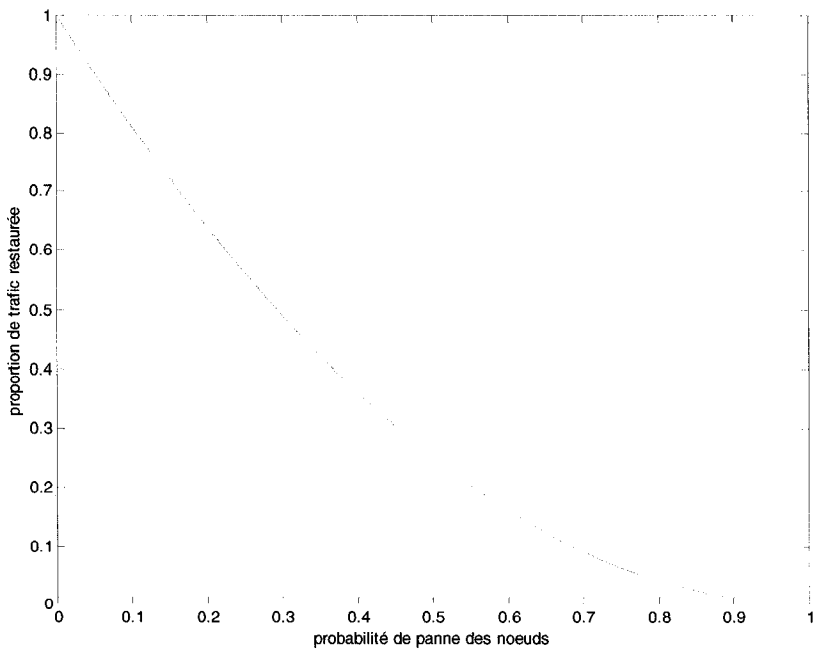

Figure 4.3 Quantité de trafic restaurée pour la solution 1 et n=2

La deuxième remarque qu'on peut déduire de ces résultats est que l'espérance de trafic restauré diminue considérablement en fonction de la longueur des chemins de secours. Cette différence est d'autant plus accentuée dans les cas de probabilité de panne médiane. Cette observation nous pousse à émettre une réserve sur l'utilité de multiplier les chemins de secours dans le cas ou cette multiplication de chemins disjoints augmenterait considérablement la longueur moyenne des chemins. Une étude plus approfondie de l'impact de cette statistique serait à suggérer dans le cas ou des développements sur les outils de simulation permettraient de réaliser cette étude de façon rigoureuse.

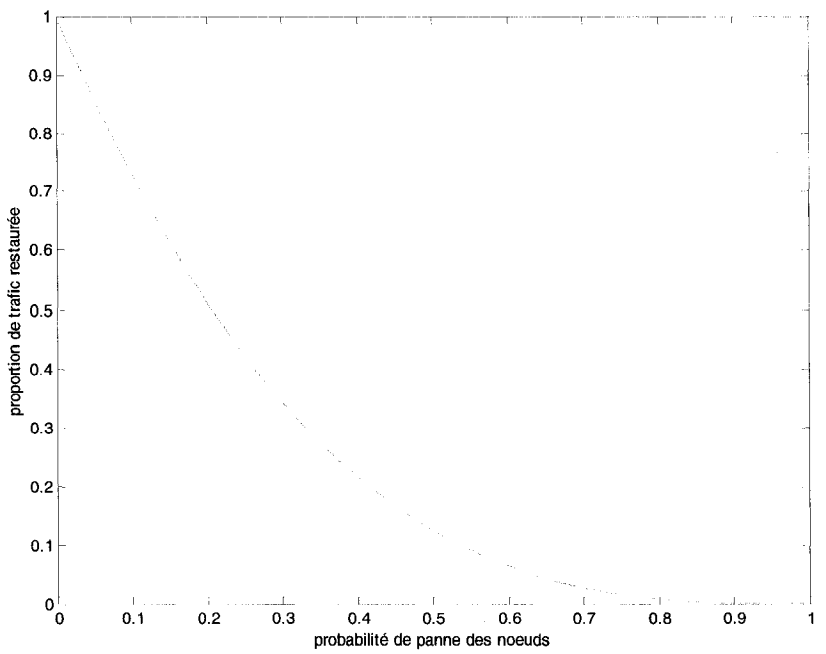

Figure 4.4 Quantité de trafic restaurée pour la solution 1 et n=3

En observant les figures de la solution 1, il serait facile de penser que la solution proposée n'a aucun mérite, puisqu'il est mathématiquement démontrée qu'elle dégrade le délai et qu'elle n'augmente pas la quantité de trafic restaurée. Toutefois, la solution 1 propose d'autres avantages qui seront détaillés dans la section probabiliste de l'étude.

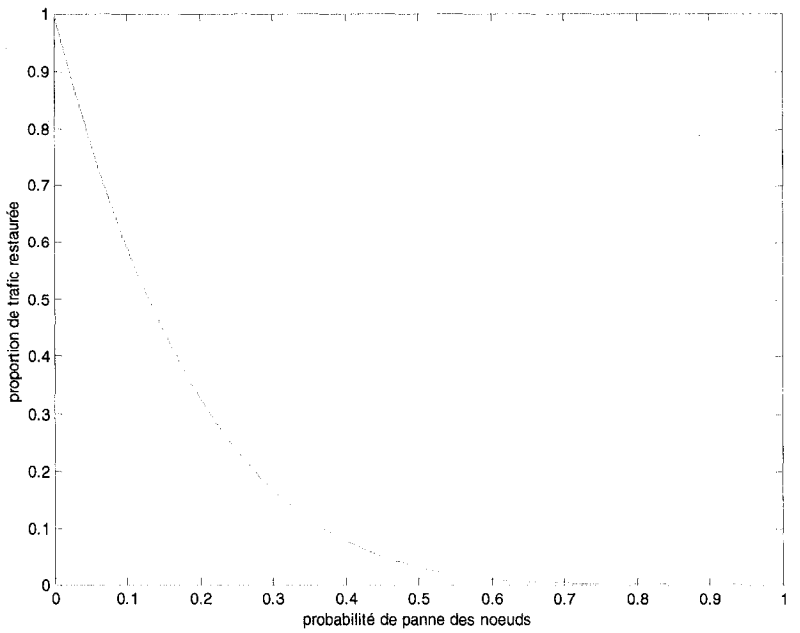

Figure 4.5 Quantité de trafic restaurée pour la solution 1 et n=5

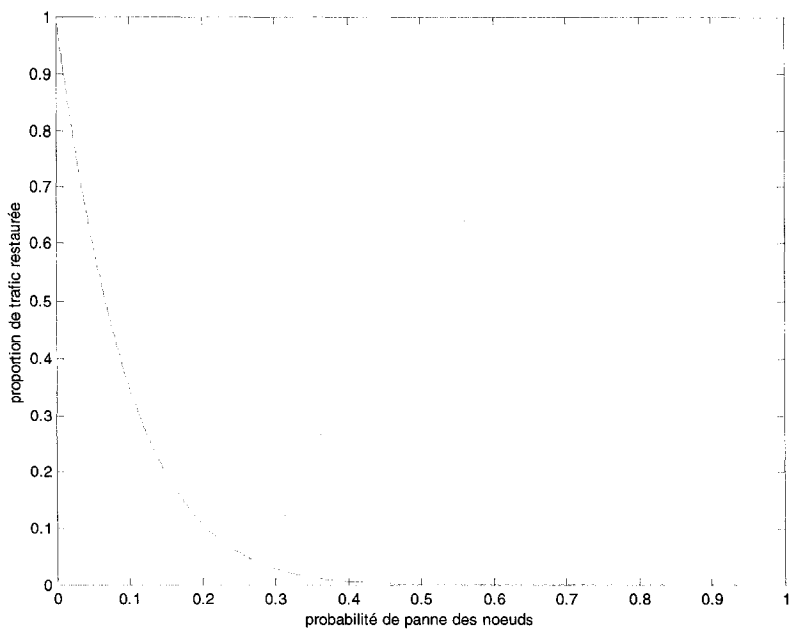

Figure 4.6 Quantité de trafic restaurée pour la solution 1 et n=10

Pour la solution 2, dont on retrouve les résultats aux figures 4.7 à 4.10, on peut remarquer que l'espérance mathématique de la quantité de trafic restaurée est largement supérieure à celle obtenue avec MPLS. Cette augmentation de l'espérance est d'autant plus marquée lorsqu'on augmente le nombre de chemins de secours. Aussi, il est possible de remarquer que la différence la plus importante se situe au niveau des probabilités de pannes médianes qui représentent la zone la plus intéressante des cas étudiés (dans la plage à faible probabilité de panne, il est souvent facile de recouvrir les pannes, tandis que pour les probabilité de panne élevées le réseau devient presque inexistant).

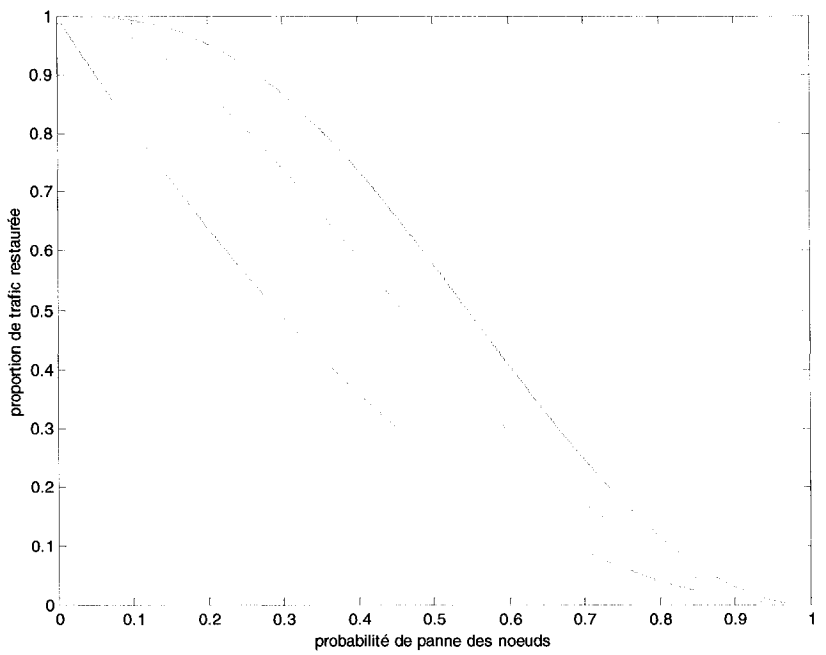

Figure 4.7 Quantité de trafic restaurée pour la solution 2 et n=2

On remarque que la solution 2 subit aussi une considérable dégradation de performance lorsque le nombre de nœuds par chemin de secours diminue. Toutefois, cette dégradation ne se fait ni plus vite, ni plus lentement que pour MPLS. Ainsi, la

solution 2 permet d'avoir une meilleure espérance mathématique de quantité de trafic restaurée que MPLS même quand $|n|$ augmente. La solution 2 est donc moins vulnérable à l'augmentation de la longueur du chemin puisqu'elle peut se permettre une légère augmentation du chemin moyen et toujours être plus performante que MPLS. Une étude approfondie de la relation entre le nombre de chemins disjoints et la variation de $|n|$ serait nécessaire pour tirer de véritables conclusions et pour déterminer précisément quel serait le choix idéal du nombre de chemins. Il est probable qu'à un certain point, l'augmentation de performance donnée par la plus grande quantité de chemins de secours soit contrebalancée par l'augmentation de |n| sur les différents chemins de secours pour obtenir un grand nombre de chemins disjoints.

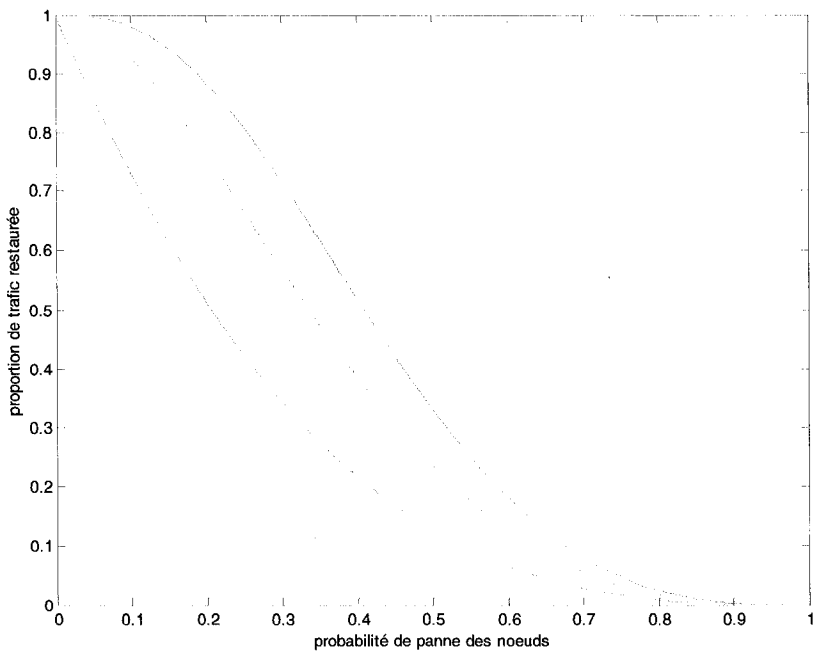

Figure 4.8 Quantité de trafic restaurée pour la solution 2 et n=3

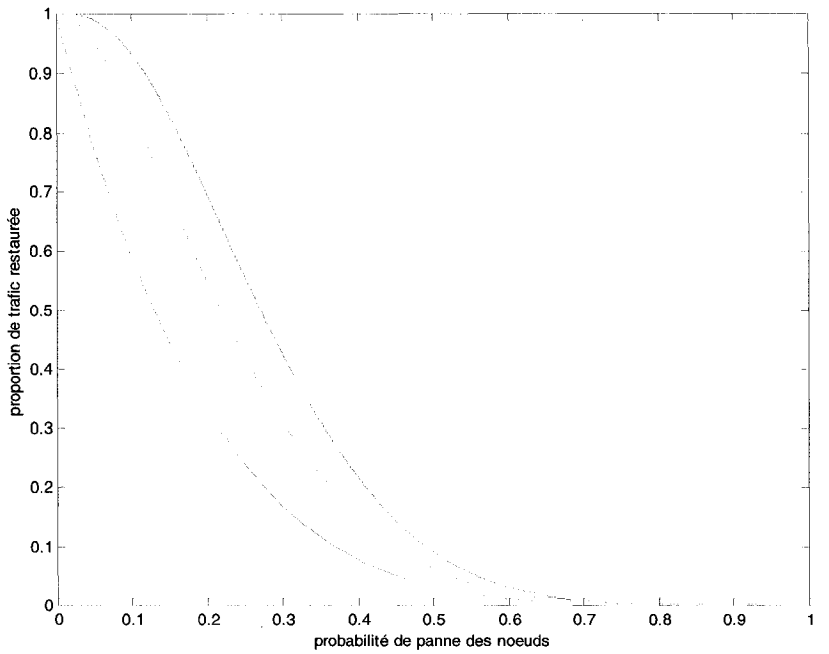

Figure 4.9 Quantité de trafic restaurée pour la solution 2 et n=5

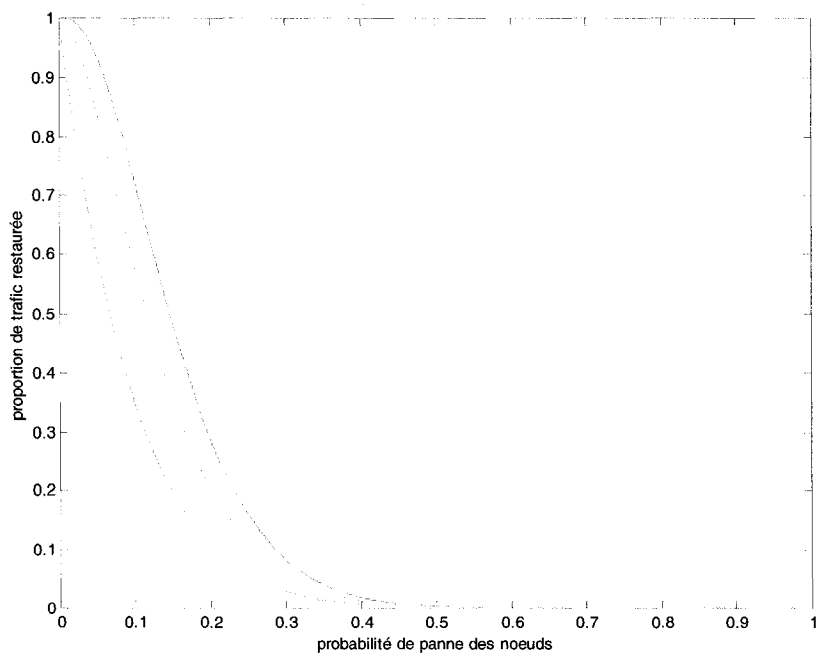

Figure 4.10 Quantité de trafic restaurée pour la solution 2 et n=10

À la lumière de ces résultats, il est tentant d'affirmer que la solution 2 est plus performante au niveau de la quantité de trafic restaurée que la protection par commutation par MPLS. L'étude du modèle probabiliste nuancera cette analyse.

4.3.2 Modèle probabiliste

Pour le modèle probabiliste, on a attribué à chaque chemin une bande passante (d'une valeur de $1/m$ pour les chemins de la solution 1, et de 1 pour les chemins de la solution 2 ainsi que pour les chemins de MPLS). On vérifie ensuite si le chemin survit à la panne en testant, pour chacun de ses $|n|$ nœuds, si le nœud subit une panne non-recouvrable. Si tous les nœuds d'un chemin ont survécu, on ajoute la valeur de bande passante attribuée préalablement à la valeur de bande passante restaurée totale. Dans le cas de la solution 2, on ajoute la valeur pour au maximum un chemin puisque l'existence de plusieurs chemins n'améliore pas la quantité de bande passante dans la solution 2.

Afin de pouvoir se comparer à la solution 1, les mêmes paramètres ont été utilisés pour la plupart des éléments. La seule différence de paramètre est le nombre de LSP à protéger qui a été augmenté à 30 pour avoir une meilleure représentation statistique pour chaque expérience. Aussi, pour générer une expérience statistique de bonne qualité, nous avons répété l'expérience 1000 fois pour chaque probabilité de panne. Les statistiques de la moyenne et de l'écart type ont été prises pour chacune de ces expériences. Les graphiques de la solution 1 seront présentés en premier, en commençant par les données sur la moyenne puis sur la variance, et les graphiques de la solution 2 suivront dans le même ordre. La légende suit la même convention que pour la méthode basée sur l'espérance.

Pour la moyenne des expériences de la solution 1, nous obtenons exactement les résultats prédits par l'espérance mathématique. Ceci améliore la confiance que nous avons envers l'analyse mathématique présentée au chapitre 3. Compte tenu de la similarité, seules les valeurs de nombre de nœuds 3, 5 et 7 ont été présentées aux Figures 4.11 à 4.13.

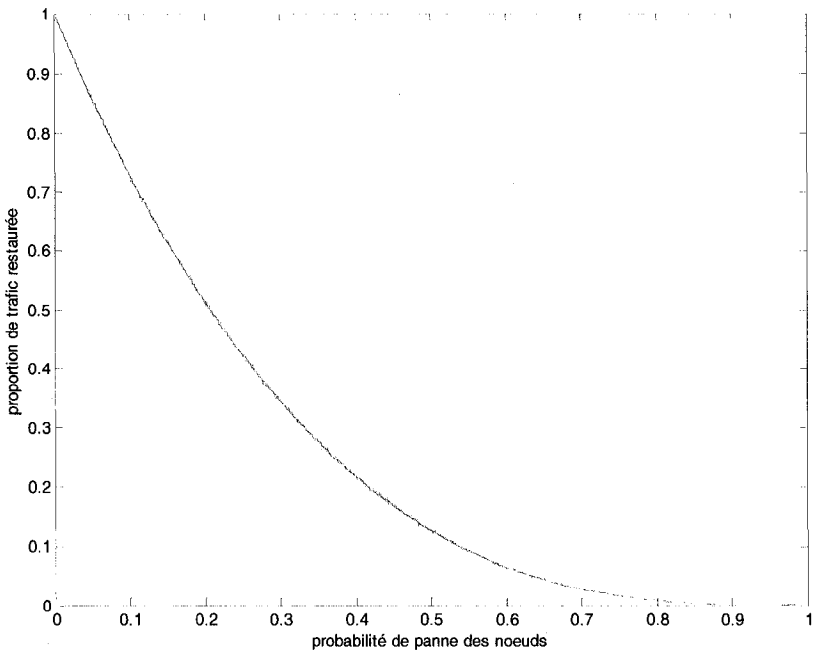

Figure 4.11 Moyenne de trafic restauré pour la solution 1 et n=3

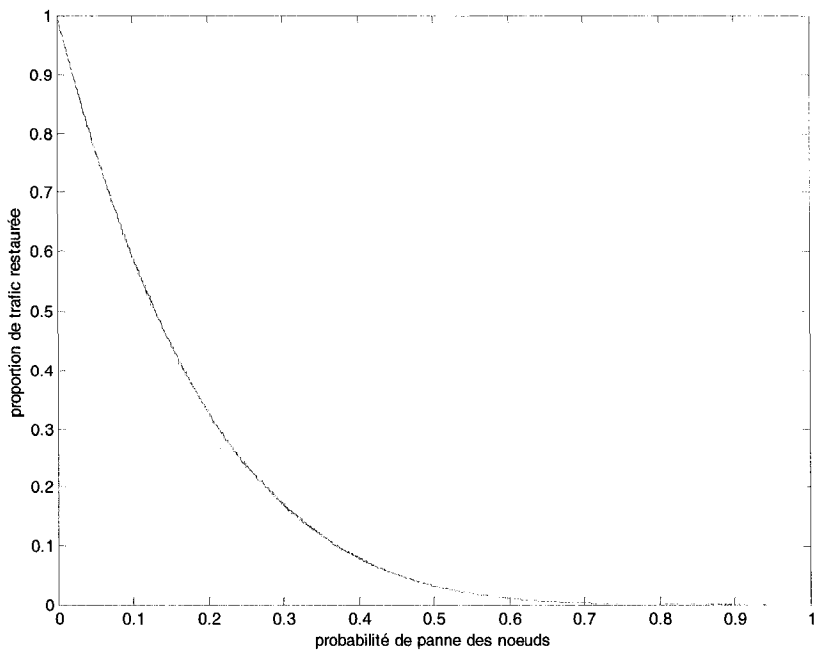

Figure 4.12 Moyenne de trafic restauré pour la solution 1 et n=5

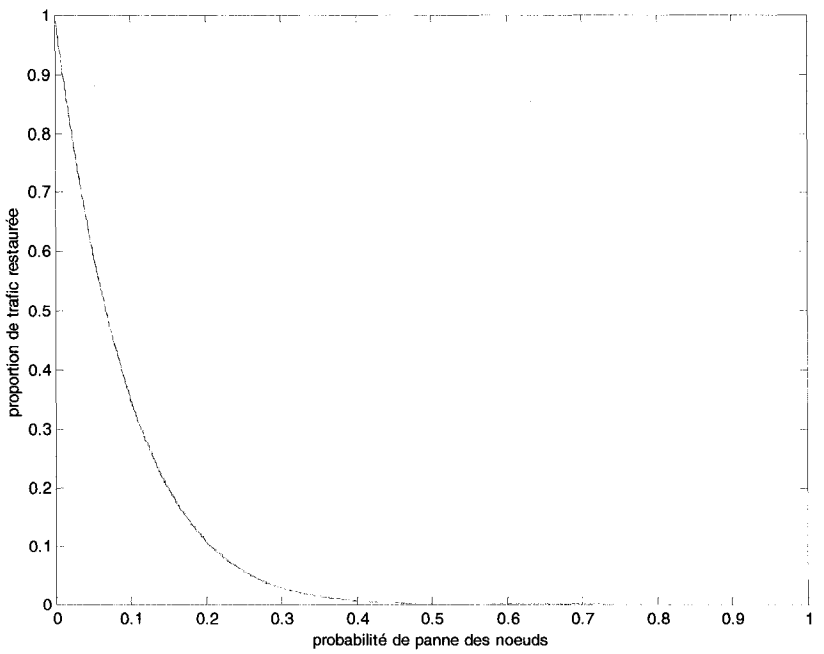

Figure 4.13 Moyenne de trafic restauré pour la solution 1 et n=10

Bien que les données sur la moyenne ne fassent que confirmer notre analyse, les données relatives à la variance permettent d'illustrer la qualité ou l'efficacité de la solution 1. On peut aisément remarquer que la variance diminue considérablement à mesure que le nombre de chemins de protection augmente. Ceci s'explique par le fait que, dans la solution 1, on obtient rarement une restauration complète, mais presque toujours une restauration partielle. On peut ainsi assurer une meilleure prédictibilité sur la disponibilité des services. Ceci correspond à ce qu'on attendait en augmentant la diversité des chemins à emprunter. Notons que cette diminution de la variance se fait malgré une distribution aléatoire des pannes non-recouvrables dans le réseau. Dans le cas où les pannes seraient concentrées géographiquement (comme dans le cas d'une attaque de déni de service), on devrait pouvoir observer des résultats encore meilleurs. Toutefois, cette conclusion devrait être validée par simulation si les avancées au niveau

des simulateurs permettaient de conduire rigoureusement cette étude. Les Figures 4.14 à 4.16 illustrent les résultats de la variance pour des valeurs de $|n|$ de 3, 5 et 10.

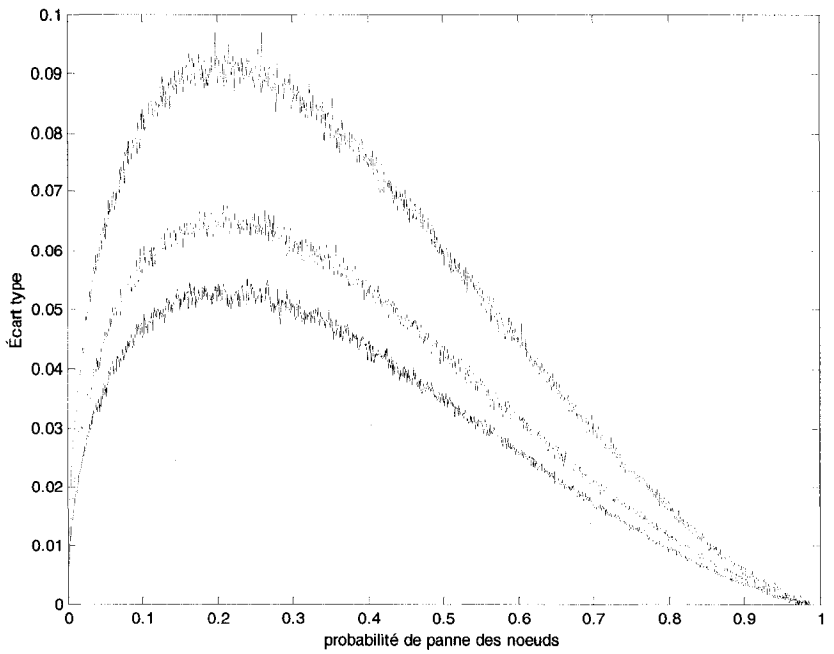

Figure 4.14 Variance du trafic restauré pour la solution 1 et n=3

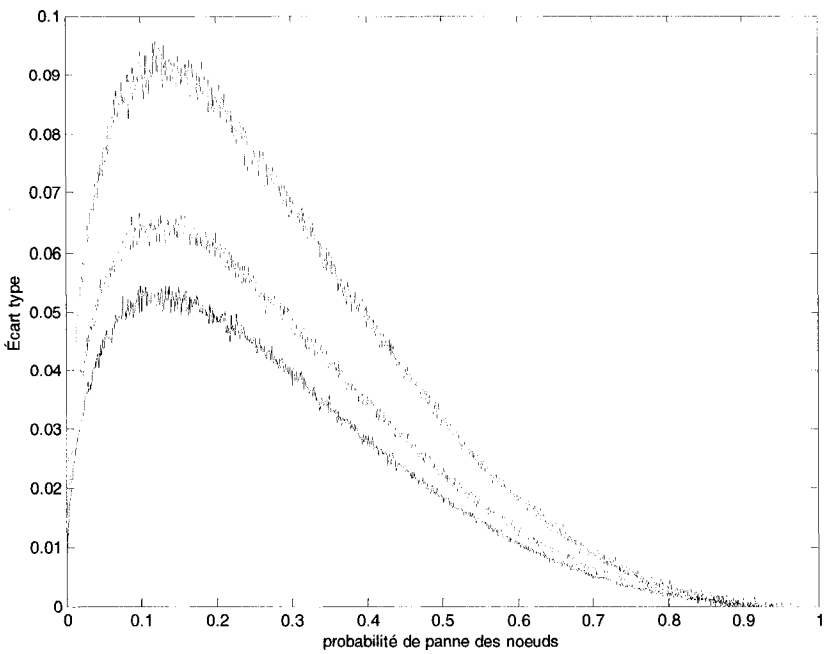

Figure 4.15 Variance du trafic restauré pour la solution 1 et n=5

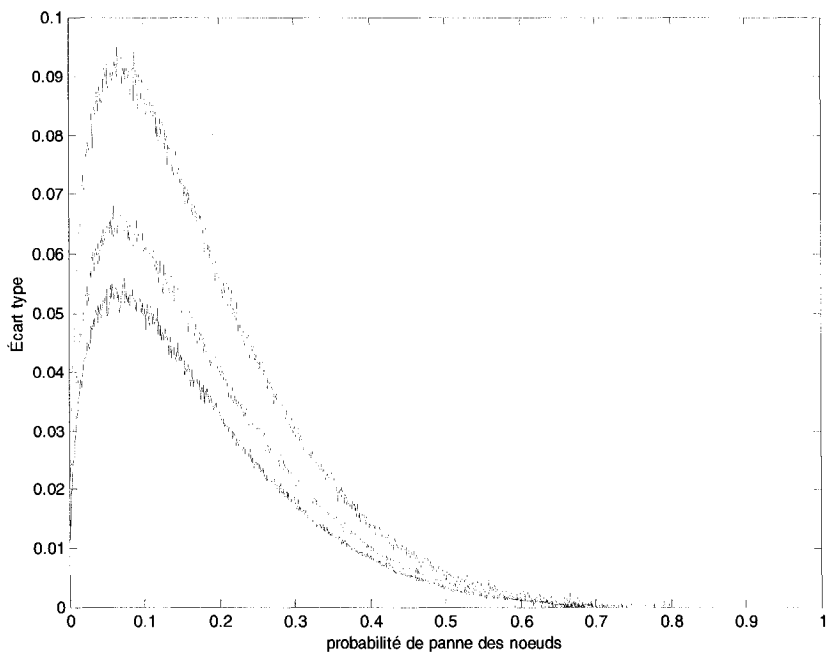

Figure 4.16 Variance du trafic restauré pour la solution 1 et n=10

Les mêmes expériences ont été réalisées pour la solution 2. Comme pour la solution 1, les valeurs moyennes sont très proches des valeurs espérées prédites au chapitre 3. Les figures 4.17 à 4.19 montrent ces résultats.

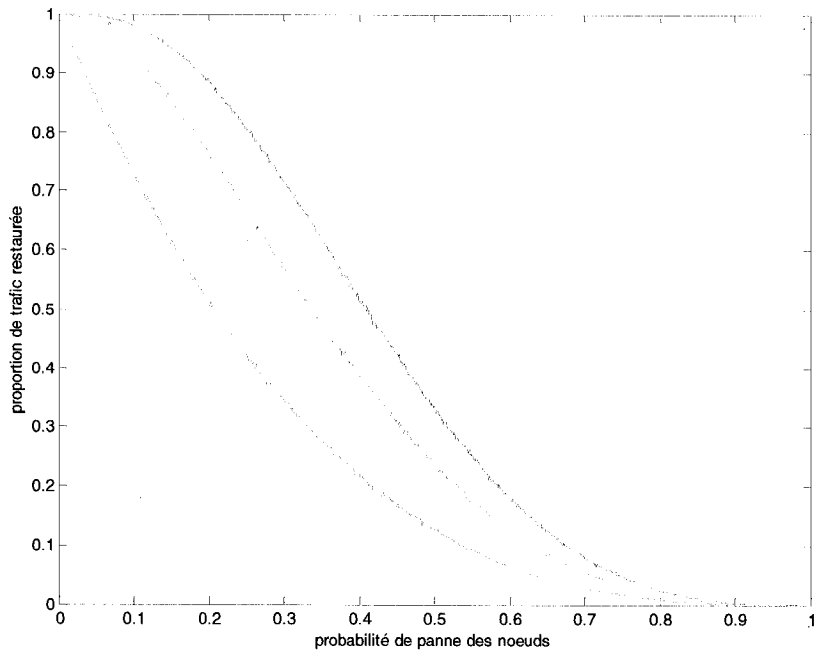

Figure 4.17 Moyenne du trafic restauré pour la solution 2 et n=3

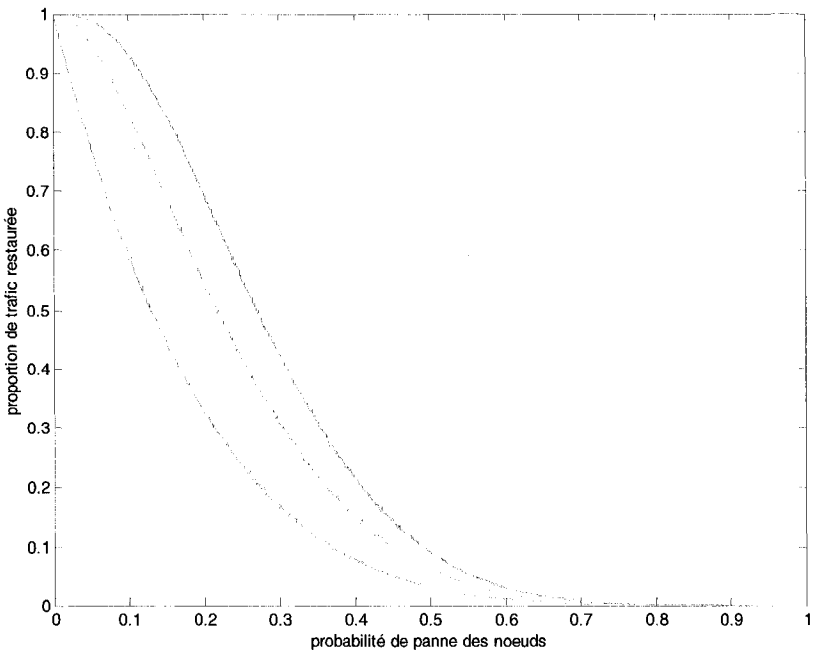

Figure 4.18 Moyenne du trafic restauré pour la solution 2 et n=5

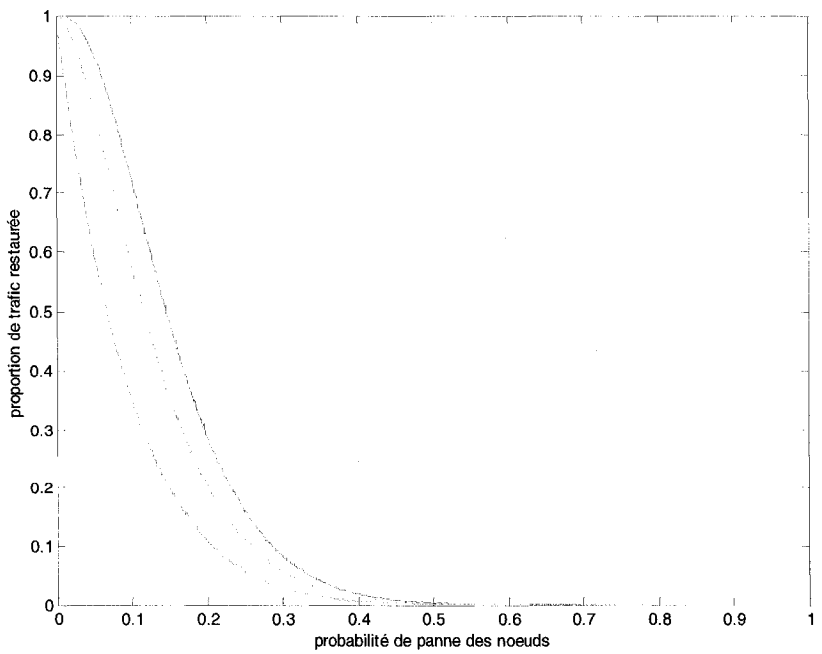

Figure 4.19 Moyenne du trafic restauré pour la solution 2 et n=10

Pour la variance, nous obtenons des résultats différents de la solution 1. Plutôt qu'une diminution de la variance, nous obtenons plutôt un déphasement de la variance vers les valeurs médianes de la probabilité de panne. On obtient les valeurs de variance maximales approximativement au point d'inflexion des courbes de la moyenne. Cette observation est donc normale puisque le point d'inflexion représente le point où on a une chance égale de survie ou de perte du chemin. Comme la solution 2 est une solution de type tout ou rien (on restaure 100% ou 0%), il est normal d'observer une grande variation. Ceci diminue la prédictibilité de la solution 2, particulièrement aux valeurs où on obtient la meilleure amélioration de l'espérance de la bande passante restaurée. Le débat porte donc sur la question de savoir s'il est préférable d'avoir une bonne prédictabilité, ou une bonne restauration (c'est-à-dire s'il est préférable d'avoir une restauration prédictible ou maximale). Les Figures 4.20 à 4.22 montrent les résultats de la variance.

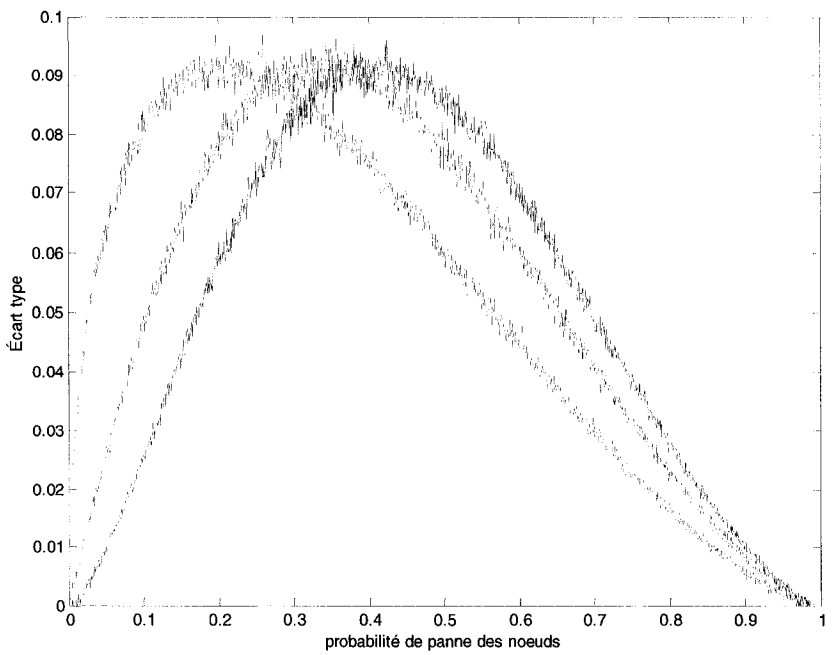

Figure 4.20 Variance du trafic restauré pour la solution 2 et n=3

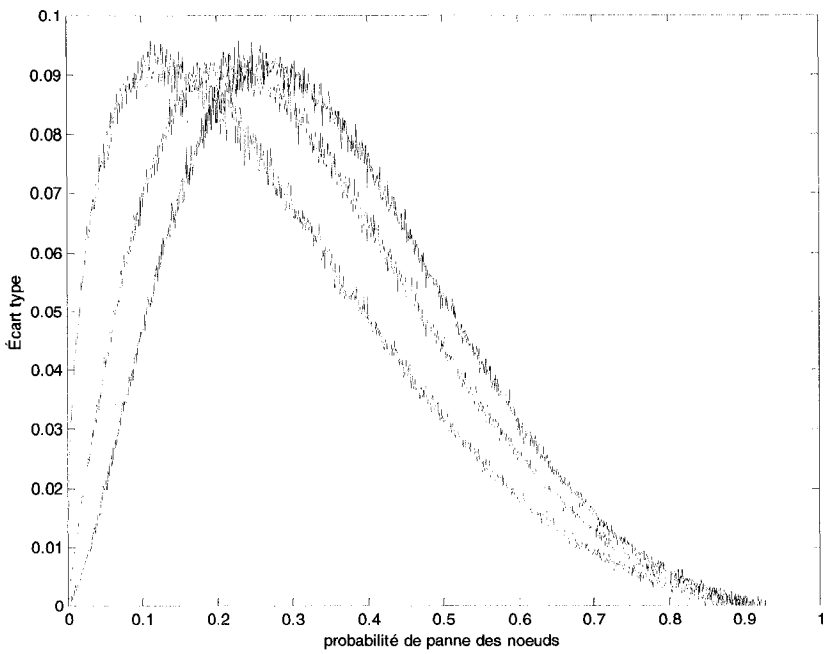

Figure 4.21 Variance du trafic restauré pour la solution 2 et n=5

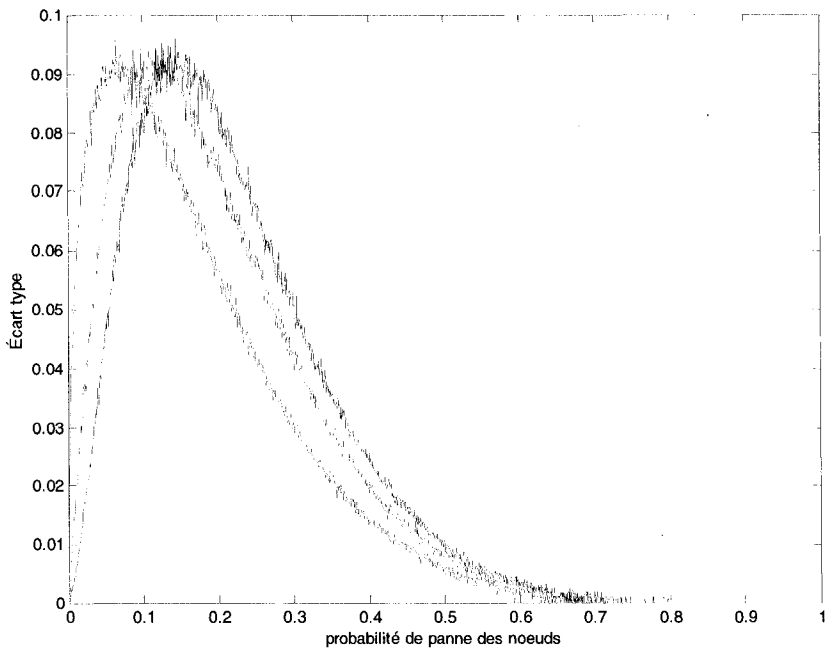

Figure 4.22 Variance du trafic restauré pour la solution 2 et n=10

À la lumière de ces résultats, nous pouvons affirmer que les solutions proposées présentent une amélioration des méthodes conventionnelles pour traiter le cas des pannes catastrophiques. La solution 1 propose une quantité de trafic restaurée équivalente, mais augmente considérablement la prédictibilité de la restauration. Cette option représente donc une restauration plus équitable. La solution 2 présente une meilleure espérance de trafic restaurée que la méthode de commutation à l'ingress par MPLS. Elle présente donc une option de restauration plus performante vis-à-vis des métriques relatives aux pannes de type catastrophique. De plus, nous avons montré que les caractéristiques de chacune des solutions (équitabilité et performance) augmentent en fonction du nombre des LSPs de protection qui sont utilisées. Ainsi, dans les réseaux avec un degré de connectivité élevé, il serait possible d'obtenir une très bonne protection.

CHAPITRE V

CONCLUSION

Ce présent mémoire avait pour objectif de présenter une architecture permettant d'assurer la survivabilité des réseaux transportant de l'information de type mission critique pour du trafic conversationnel et meilleur effort. Nous allons en premier lieu rappeler les points qui confirment la réalisation de cet objectif, puis nous allons discuter des limitations des approches proposées et terminer en proposant quelques pistes de recherches futures.

5.1 Synthèse des travaux

Commençons par rappeler que, depuis le 11 septembre 2001, la communauté œuvrant en réseautique a été mise au courant de façon brutale de la sévérité des effets d'une panne catastrophique. Les effets d'une telle panne ont été jugés trop grands pour que la faible probabilité d'occurrence rende le risque acceptable. De plus, avec l'avènement des réseaux de prochaine génération, en particulier les réseaux de type tout-IP, on pourra observer des réseaux transportant plusieurs types de services. Pour certains de ces services, une certaine qualité de service doit être garantie. Ceci complexifie grandement le problème puisque les objectifs de qualité de service et les objectifs de disponibilités des ressources (rendue d'autant plus critique en situation de panne) sont des objectifs concurrents. L'ensemble de ces facteurs fait du domaine de la survivabilité vis-à-vis des pannes catastrophiques un domaine riche en perspectives scientifiques.

Toutefois, en dépit de l'attrait scientifique du problème, la revue de littérature a montré que peu de solutions existaient pour ce problème. On retrouvait un grand nombre de solutions pour assurer la qualité du service malgré une panne de réseau (mentionnons entre autres la protection par commutation et le FAST-REROUTE de MPLS, ou encore

le APS de SONET). Malheureusement, ces solutions sont souvent adaptées pour réagir seulement à des pannes uniques. Aussi, ces solutions consomment beaucoup de ressources réseau pour assurer la rapidité de la commutation. Elles sont donc mal adaptées aux situations de pannes catastrophiques, où les ressources réseau sont rares et où les pannes multiples sont la norme. De l'autre côté de la médaille, on retrouve les solutions adaptées aux pannes catastrophiques. Ces solutions, issues de la crainte des attaques de missiles pendant la guerre froide, permettent d'assurer la connectivité maximale du réseau tant que cette connectivité est possible. Malheureusement, avec l'inclusion du trafic ayant des requis élevés de qualité de service, ces méthodes ont mal vieilli. Elles ne peuvent tout simplement pas fournir les délais de restauration nécessaire pour assurer la continuité des services. On ne retrouve donc pas dans la littérature une solution permettant de remplir l'objectif de ce présent mémoire.

Pour combler cette lacune, nous avons proposé deux architectures pour assurer la protection de réseaux de type métropolitains supportant MPLS. Dans la première architecture, le trafic possédant des requis stricts de qualité de service se voyait attribué m chemins de protection de bande passante b/m plutôt qu'un seul chemin de protection de bande passante b. En cas de panne, le trafic est réparti sur les chemins de secours ayant survécu. Une fois la restauration du trafic à requis strict de qualité de service achevé, le reste du trafic est restauré par les méthodes de la couche 3. L'objectif de cette solution est de sacrifier une partie de la qualité de service pour assurer la préservation de la connexion. Dans la deuxième méthode, plutôt que de répartir le trafic sur les chemins ayant survécu, le trafic est envoyé, accompagné d'une renégociation de la bande passante. On commute le trafic sur les différents chemins de protection jusqu'à ce qu'on trouve un chemin acceptant la demande. Ceci permet de conserver la même qualité de service qu'avant la panne. Toutefois, on doit sacrifier le temps de restauration pour obtenir cette augmentation de performance.

Finalement, afin de prouver les assertions avancées dans le chapitre 3, nous avons effectué une évaluation de performance des solutions proposées. Cette analyse de performance comparait les solutions avec la méthode de protection par commutation à

l'ingress offerte par MPLS. Une preuve de fonctionnement a d'abord été réalisée à l'aide du logiciel OPNET. Dans cette preuve de fonctionnement, nous avons pu observer que la moins performante des solutions proposées permettait d'égaler la quantité de trafic restaurée offerte par MPLS. Ceci correspondait au résultat prédit dans le chapitre 3. Pour généraliser cette étude, ce mémoire a présenté les résultats pour différentes valeurs de probabilité de panne, de nombre de nœuds par chemin de protection et de LSPs de protection pour les deux solutions avancées. Aussi, des études statistiques validant ces valeurs ont été montrées avec des statistiques sur la variance des résultats sur 1000 expériences. Dans ces expériences, on a pu voir que la performance, en fonction de la métrique du trafic restaurée, de la solution 2 était supérieure à celle de MPLS dans tous les cas. Aussi, nous avons vu que la solution 1 offrait une méthode également performante au niveau de la quantité de trafic restaurée, mais plus prédictible que la méthode proposée par MPLS. On peut donc affirmer que les deux solutions présentent une amélioration vis-à-vis des solutions existantes pour la survivabilité des systèmes d'informations face aux pannes catastrophiques.

5.2 Limitations des travaux

Les solutions proposées possèdent malgré tout un certain nombre de limitations. La limitation la plus importante est le manque de données concernant l'évaluation de performance. Toute l'évaluation de performance se base sur le modèle analytique présenté dans le chapitre 3 qui n'a été validé que dans un cas particulier. Malheureusement, l'état actuel de développement des simulateurs de réseau ne permet pas de réaliser cette analyse de performance de façon optimale. Une analyse de performance se basant sur une implémentation d'un prototype serait aussi beaucoup trop coûteuse pour être envisageable. Afin de palier à cette limitation, le plan d'expérience d'une analyse de performance dans les règles a été présenté au chapitre 4.

Une deuxième limitation importante est la nécessité d'avoir un réseau incluant une couche 2,5 (MPLS ou ATM). Ces deux technologies ont, bien sûr, atteint un taux de

pénétration assez élevé, mais ce taux n'est pas encore 100%. Les architectures proposées ne tiennent donc pas compte de l'intégration des réseaux ne supportant pas ces technologies. Les solutions ne tiennent pas compte non plus des mécanismes de protection pouvant être déployées sur les autres couches du modèle OSI qui pourraient être plus performantes ou avoir des interactions avec les solutions présentées. Toutefois, les solutions proposées n'empêchent nullement le déploiement d'architectures de protections additionnels basés sur d'autres architecture réseau ou logicielles.

Une troisième limitation est une limitation au niveau de la topologie. En posant l'hypothèse d'un réseau métropolitain, il est devenu possible d'avoir une topologie réseau possédant un degré assez élevé. Ce degré élevé permettait de bénéficier de réseaux possédant une grande diversité. Cette diversité est nécessaire au déploiement de chemin de secours (c'est-à-dire des chemins disjoints de nœud) multiples. Ce type de topologie est peu commun si on pousse l'étude vers l'accès, qui est généralement configuré en étoile.

Hormis les limitations mentionnées, les solutions proposées possèdent sûrement de multiples autres limitations qui restent à découvrir. En effet, compte tenu de la nouveauté du champ d'étude traité dans ce mémoire, il est prématuré de proposer une solution complète, ou même de prétendre proposer une solution optimale. L'auteur espère donc que les limitations qui ne sont pas mentionnées seront investiguées par les prochaines générations de chercheurs.

5.3 Indications de recherche future

Plusieurs pistes de recherche sont fertiles à de nouveaux développements. Tout d'abord, une analyse de performance approfondie de la survivabilité d'un réseau de prochaine génération à une panne catastrophique permettrait de quantifier précisment les développements dans le domaine. Aussi, le développement de méthodes de protection à d'autres niveaux de la couche OSI (physique, ou applicatives par exemple) ainsi que le développement de méthodes incorporant d'autres technologies (UMTS, GMPLS, etc.)

permettraient d'enrichir le champ d'étude. Le développement de méthodes de protection contre les pannes catastrophiques applicables à des réseaux plus diversifiés que les réseaux métropolitains seraient une autre avenue profitable.

Une autre piste de généralisation serait la généralisation de l'implémentation. En effet, puisque ATM offre (au moins) les mêmes fonctionnalités que MPLS, les propositions pourraient aussi être implémentées à l'aide d'ATM. OSPF aussi pourrait être remplacé par d'autres technologies (IS-IS par exemple). De plus, il est possible d'utiliser la même architecture pour développer une méthode de protection locale plutôt que globale pour améliore l'implémentation. Toutes ces pistes, qui tiennent plus du développement que de la recherche, pourraient bénéficier de travaux futurs.

Dans une perspective plus large, plusieurs champs d'études connexes pourraient apporter d'importantes contributions au champ de la survivabilité aux pannes catastrophiques. Mentionnons principalement le développement des réseaux intelligents et autonomes qui peuvent fonctionner malgré le partitionnement. Ce champ d'étude, principalement mis de l'avant par la communauté étudiant les agents mobiles, pourrait offrir des solutions concrètes pour survivre à des pannes causant des partitionnements. Bien que la connectivité de bout en bout soit toujours impossible, la possibilité de fonctionner localement serait déjà une amélioration majeure.

Un autre champ d'étude très prometteur pour la survivabilité aux pannes catastrophiques est celui concernant les réseaux mobiles ad hoc. Ces réseaux sans configuration fixe ne possèdent pas d'infrastructure vulnérable aux catastrophes. Ceci permettrait de diminuer considérablement la sévérité des pannes. De plus, à l'instar des réseaux subissant des pannes catastrophiques, la topologie des réseaux ad hoc varie de façon rapide et dynamique. Bien que les causes de cette variation dynamique de la topologie (la perte d'éléments réseau pour le cas des pannes et la mobilité et le manque de fiabilité pour les réseaux mobiles ad hoc) soient très différentes, les avancées au niveau de l'administration de réseaux dynamiquement reconfigurables pourraient inspirer des méthodes de reconfiguration rapide des réseaux filaires subissant des pannes catastrophiques. Avec des mécanismes efficaces d'administration répartie et de

reconfiguration dynamique du réseau, il sera peut-être possible, un jour, d'avoir des réseau dont la capacité d'auto-regénération sera capable d'accommoder les pannes catastrophiques.

BIBLIOGRAPHIE

[1] M. Ajtai, N. Alon, J. Bruck, R. Cypher, C. T. Ho, M. Naor and E. Szemerèdi, "Fault tolerant graphs, perfect hash functions and disjoint paths", *Proc. 33 IEEE FOCS*, Pittsburgh, IEEE (1992), 693-702.

[2] A. Autenrieth, A. Kirstädter, "Engineering End-to-End IP Resilience Using Resilience-Differentiated QoS", *IEEE Communications Magazine*, Vol. 40, No. 1, Jan 2002, pp. 50-57.

[3] A. Autenrieth, A. Kirstädter, "RD-QoS - The Integrated Provisioning of Resilience and QoS in MPLS-Based Networks", *IEEE International Conference on Communications (ICC 2002)*, New York, USA, April 28 - May 02, 2002, pp. 1174-1178.

[4] A. Autenrieth, A. Kirstädter "Fault Tolerance and Resilience Issues in IP based Networks", *Second International Workshop on the Design of Reliable Communication Networks*, München, April 10-12, 2000.

[5] A. Kirstädter, A. Autenrieth, *An Extended QoS Architecture Supporting Differentiated Resilience Requirements of IP Services*, IETF Draft, draft-kirstaedter-extqosarch-00.txt, August 9, 2000.

[6] A. Autenrieth, A. Kirstädter, "Provisioning of Differentiated IP Resilience and QoS-An Integrated Approach", *ITG Workshop "IP in Telekommunikationsnetzen"*, Bremen, January 26, 2001.

[7] A. Autenrieth, A. Kirstädter, "Components of MPLS Recovery Supporting Differentiated Resilience Requirements", *7th EUNICE 2001*, September 3-5, 2001, Paris, France.

[8] A. Autenrieth, A. Kirstädter, "Components of MPLS Recovery Supporting Differentiated Resilience Requirements", *IFIP Workshop on IP and ATM Traffic Management, WATM 2001,* September 3-5, 2001, Paris, France.

[9] D. Awduche, L. Berger, T.Li, V.Srinivasan, G.Swallow, *RSVP-TE: Extension to RSVP for LSP tunnels*, RFC3029, December 2001.

[10] Baran, "On distributed communications", Memorandum, *Rand corporation*, RM-3420-PR, August 1964.

[11] R. Braden, L. Zhang, S. Berson, S. Herzog S. Jamin, *Resource Reservation protocol (RSVP) -- Version 1 Functional specification*, RFC2205, Sept 1997.

[12] R. Boutaba et Andreas Polyrakis, "Projecting Advanced Enterprise Network and Services Management to Active Networks", *IEEE Network*, January/February 2002.

[13] M. Clouqueur, W. Grover, D. Leung et O. Shai., "Mining the Rings : Strategies for Ring-to-Mesh Evolution", *3rd international workshop on the design of reliable communication networks (DRCN 2001)*, Budapest, Hungary, October 2001.

[14] G. Dorvius, "Routage proactif alterné basé sur la qualité de service UMTS ", mémoire de maîtrise présenté à l'Université de Montréal, École Polytechnique, Août 2003.

[15] M. Elder, Fault Tolerance in Critical Information Systems, these de doctorat présenté à University of Virginia, May 2001.

[16] G. Ellinas, A. G. Hailemariam et T. Stern , "Protection cycles in mesh WDM networks", *IEEE journal on selected areas of communication*, Vol 18, No 10, October 2000.

[17] A. Farrel, B. Miller, *Surviving Failures in MPLS Networks*, February 2001, Data Connection Limited, Enfield, UK.

[18] W. Grover, "Case Study of Survivable Ring, Mesh and Mesh-Arc Design", *IEEE Global Telecommunications Conference*, December 1992.

[19] C. Hang, V. Sharma, S. Makam, K. Owens, *A Path Protection/Restoration Mechanism for MPLS Networks*, (work in progress) Internet Draft draft-chang-mpls-path-protection, July 2000.

[20] D. Haskin, R. Krishnan, *A Method for Setting an Alternative Label Switched Paths to Handle Fast Reroute*, (work in progress) Internet Draft draft-haskin-mpls-fast-reroute, November 2000.

[21] IEC (International Engineering Consortium), "Distributed Network Intelligence Tutorial", http://www.iec.org/online/tutorials/dist_net/

[22] IETF Network Working Group, "rfc3386 : Network Hierarchy and Multilayer Survivability", November 2002.

[23] S. Kini, M. Kodialam, T.V. Lakshman, C. Villamizar, *Shared backup Label Switched Path restoration*, (work in progress) Internet Draft draft-kini-restoration-shared-backup, October 2000.

[24] R. Krishnan, D. Haskin, *Extension to RSVP to Handle Establishment of Alternate Label-Switched-Paths for Fast Reroute*, (work in progress), Internet Draft draft-krishnan-mpls-reroute-resvpext-00.txt, June 1999.

[25] S. Pierre, "A Tabu-Search Aprroach for Designing Computer-Network Topologies with Unreliable Components", *IEEE transactions on reliability*, Vol. 46, No 3, September 1997.

[26] MATLAB, fichier d'aide de la function *rand( )* (help rand)

[27] B. Sanso et F. Soumis, "Communication & transportation network reliability using routing models", IEEE transactions on reliability, Vol. 40, No 1, April 1991.

[28] P. Stavroulakis, "Reliability, survivability and quality of large scale telecommunication systems; Case study : Olympic Games", Éditions John Wiley & Sons, 2003.

[29] D. Tennenhouse, J. Smith, W. D. Sincoskie, D. Wetherall et G. Minden, "A Survey of Active Network Research", *IEEE Communications Magazine*, January 1997.

[30] Ernesto Q. Vieira Martins, M. Margarida, B. Pascoal et J. L. E. Santos, "The K shortest paths problem", Research Report, *CISUC*, June 1998.

[31] E. Q. Vieira Martins, M. Margarida, B. Pascoal et J. L. E. Santos, "A New Improvement for a *K* Shortest Paths Algorithm", *Investigação Operacional*, 21:47-60, 2001.

[32] R. S. Wilkov, "Analysis and design of reliable computer networks", Communications, IEEE transactions on [legacy, pre-988], Vol. 20, No. 3, June 1972, pp. 660-678.

ANNEXE 1

ARCHITECTURE DE PROTECTION SONET

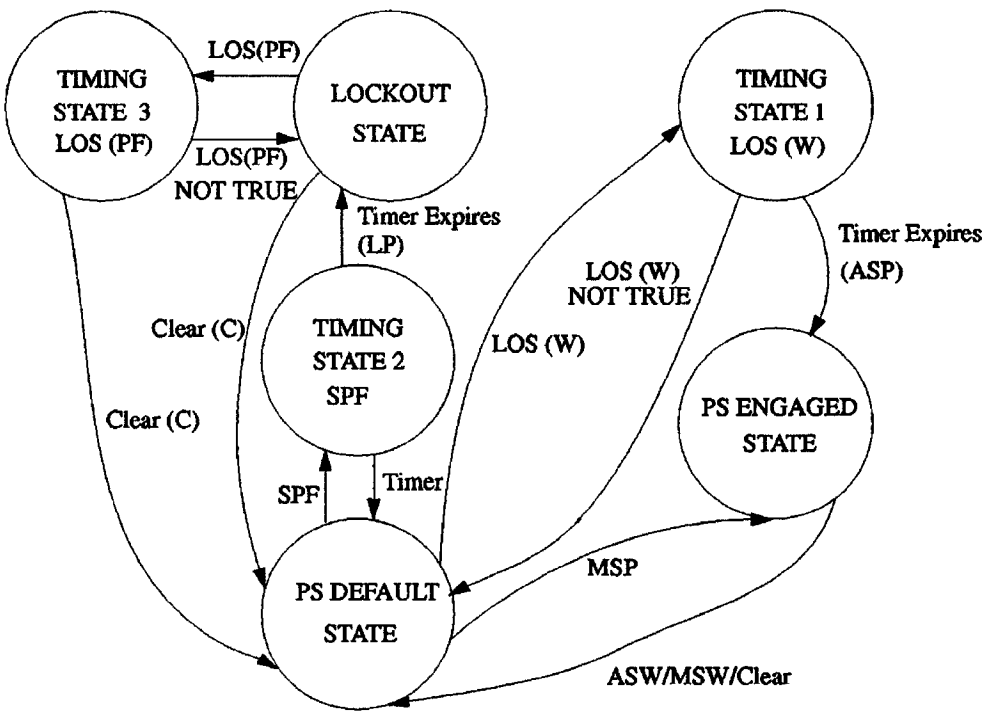

Diagramme d'état de la protection par commutation [16]

ANNEXE 2

ARCHITECTURE DE PROTECTION IP

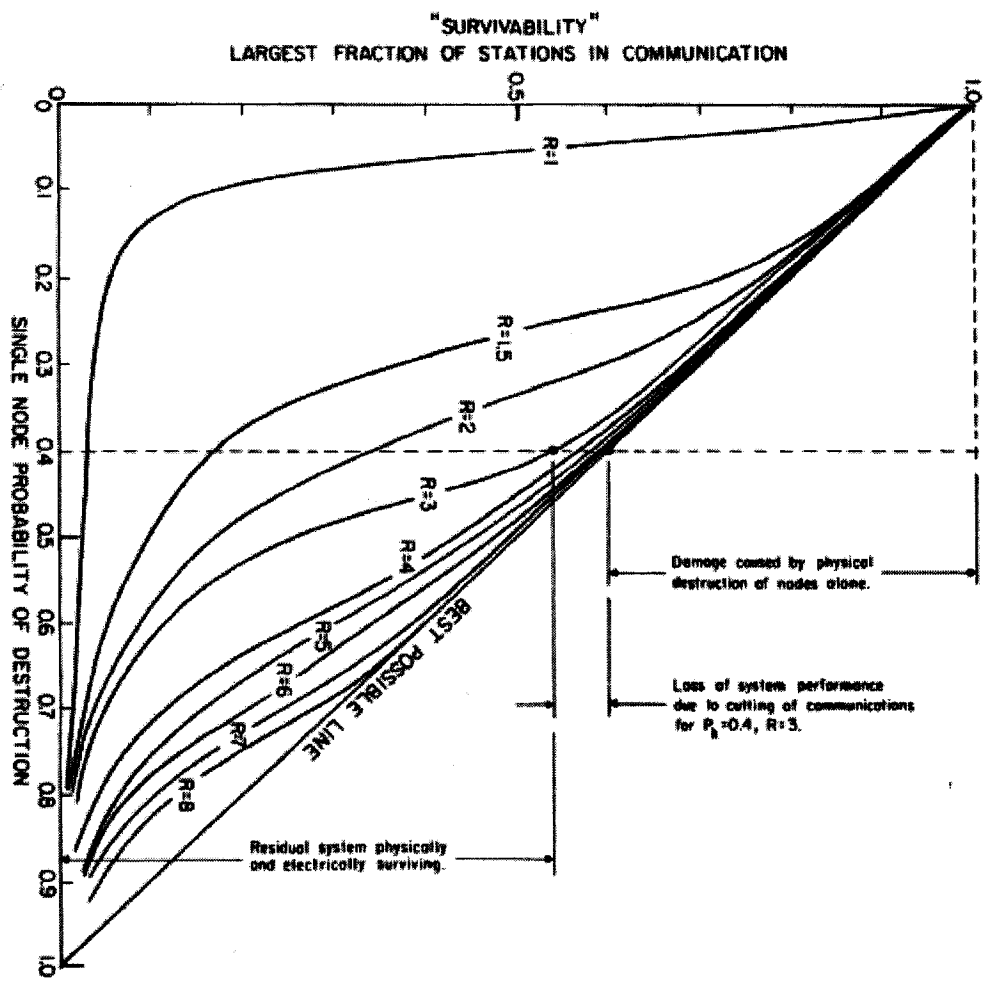

FIG. 4 — Perfect Switching in a Distributed Network — Sensitivity to Node Destruction, 100% of Links Operative.

Mesure de la survivabilité d'un réseau IP à la destruction de nœuds [10]

ANNEXE 3

TOPOLOGIE DU RÉSEAU DE TEST